IECO

Volume 6 No 3 Year 2023

■ Volume 6
■ No 3
■ Year 2023

University of Sistan and Baluchestan

Iranian Association of Electrical and Electronics Engineers

International Journal Of

Industrial Electronics Control and Optimization

In This Issue:

Research Articles:

IECO

International Journal Of
Industrial Electronics Control and Optimization

About Journal

The University of Sistan and Baluchestan entered into strategic partnership with Iranian Association of Electrical and Electronic Engineers (IAEEE) to publish the **International Journal of Industrial Electronics Control and Optimization (IECO)**. The IECO is a refereed international journal which presents to the international scientific community important results of work in these fields, whether in the form of modeling simulation, analysis, fundamental research, development, application, design or real-time implementation. The scope of IECO is broad, encompassing all aspects of Industrial Electronics, Control and Optimization.

Note: International Journal of Industrial Electronics, Control and Optimization (IECO) has qualified to **ACADEMIC RESEARCH JOURNAL (ELMI-PAJOHESHI)** status certified by the ministry of Science, Research and Technology of Iran (No. 231566/3/18 dated 1396/10/09), and is published by the University of Sistan and Baluchestan through a formal partnership (No. 952/2/1500 dated 1395/11/04) with Iranian Association of Electrical and electronic Engineers (IAEEE) in order to develop scientific and research cooperation.

Aims and Scope

International Journal of Industrial Electronics, Control and Optimization (IECO) is a Peer reviewed journal of advanced and state-of-the-art in the science and engineering of Industrial Electronics, Control and Optimization. Its Scope encompasses the applications of Industrial Electronics, power systems, control, optimization and computational intelligence for the enhancement of industrial and manufacturing system and processes. The scope of the journal include the following:

I. Industrial Electronics

- Low and high-power converters
- Renewable energy
- Drive control techniques
- Techniques for advanced power semiconductor devices
- Power quality and utility applications
- Communications
- Flexible AC Transmission Systems (FACTS)
- Control in power electronics
- Electromagnetic and thermal performance of electronic power converters
- Motion control, robotics, sensors and actuators
- Fault detection and diagnosis
- Power systems
- Factory automation, communication, and computer networks

II. Control

- Adaptive control
- Control of process systems
- Control theory
- Data processing
- Design of control systems
- Hybrid systems
- Identification and observation
- Intelligent systems
- Model-predictive control
- Optimal control

- Robust control
- Fractional order systems

III. Optimization

- Ant Colony
- Chaos Theory
- Evolutionary Computing
- Fuzzy Computing
- Hybrid Methods
- Immunological Computing
- Neuro Computing
- Particle Swarm
- Probabilistic Computing
- Rough Sets
- Wavelet

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Using Improved DDAO Algorithm to Solve Economic Emission Load Dispatch Problem in Presence of Wind Farms

Mehdi Shafiee[1] | Abbas-Ali Zamani[2] | Mehdi Sajadinia[3]

Department of Electrical Engineering, Technical and Vocational University (TVU), Tehran, Iran, Iran.[1,2,3]
Corresponding author's email: m-shafiee@tvu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 2023-March-24
Received in revised form: 2023-June-09
Accepted: 2023-July-09
Published online: 2023-July-24

Keywords:
Economic emission,
Load dispatch,
Wind farm,
Optimization algorithm.

ABSTRACT

In power systems planning, economic load dispatch considering the uncertainty of renewable energy sources is one of the most important challenges that researchers have been concerned about. Complex operational constraints, non-convex cost functions of power generation, and some uncertainties make it difficult to solve this problem through conventional optimization techniques. In this article, an improved dynamic differential annealed optimization (IDDAO) meta-heuristic algorithm, which is an improved version of the dynamic differential annealed optimization (DDAO) algorithm has been introduced. This algorithm has been used to solve the economic emission load dispatch (EELD) problem in power systems that include wind farms, and the performance of the proposed technique was evaluated in the IEEE 40-unit and 6-unit standard test systems. The results obtained from numerical simulations demonstrate the profound accuracy and convergence speed of the proposed IDDAO algorithm compared to conventional optimization algorithms including, PSO, GSA, and DDAO, while independent runs indicate the robustness and stability of the proposed algorithm.

I.Introduction

Due to the ever-increasing progress of industries and the expansion of cities, the need for electric energy and consequently, thermal power generation has significantly increased. In recent years, the emission of harmful gases such as sulfur oxide and nitrogen oxide, which have caused atmospheric pollution and intensified global warming, has become a critical issue. One of the ways to limit the emission of these gases is to apply stricter policies on thermal units [1, 2]. To apply new regulations and tax considerations for excessive greenhouse gas emissions, a problem combining load economic dispatch and emission constraints known as the load economic emission dispatch problem is introduced [2].

Economic load dispatch has an important and sensitive task in the operation and planning of power systems in such a way that it requires the generation units to meet the load demand while the total power generation cost is minimized and various technical and operational limitations of this process are met [2, 3].

One of the most important solutions to reduce environmental pollutants is to replace fossil fuels with renewable energy, and one of the most important and cost-effective sources of renewable energy is wind energy [4]. The use of wind energy in the generation of electrical energy does not cause environmental pollution, and its maintenance cost is reasonable compared to other renewable energy sources. However, due to the random nature of wind speed forecasting,

solving the problem of economic load dispatch with the participation of wind units has many challenges [5, 6].

Classical optimization methods such as gradient descent, Landa's iteration method, linear programming method and dynamic programming in the simple economic load dispatch problem provide suitable results when the cost function is smooth, continuous, and differentiable [7]. These methods don't provide suitable answers when the cost function is non-convex, discontinuous, and non-derivative and there are various electrical and mechanical constraints such as the effect of the input valve, generation and consumption balance constraints, generation limits, prohibited zones, the limit of the permitted rate of change of power and the losses of the transmission network. These problems have led researchers to use intuitive optimization algorithms to solve this problem [8-13].

In recent years, different optimization algorithms have been used to solve the problem of economic load dispatch by considering the effect of wind units, including the PSO optimization algorithm and its improved variants called EMOPSO [10], or the gravitational acceleration enhanced particle swarm optimization (GAEPSO) algorithm [14], CMOPEO-EED algorithm [15], EMA algorithm [16], NSGA-II algorithm [17], CSCA algorithm [18], IMOBSO algorithm [19], NSGWO algorithm [20], MQLSA algorithm [21], MODE algorithm [22], WMA algorithm [23], COOT algorithm [24] and CRO algorithm [25]. In the following, some of the articles presented in this regard are mentioned.

In [26], a hybrid method combined with linear and non-linear programming is employed to tackle with non-convexity of economic dispatch, considering the valve point effects. The woodpecker mating algorithm (WMA) is used by [23] to solve the economic dispatch problem considering the nonlinear properties of generators. In the presence of RESs, the chaotic slap swarm optimization algorithm has been applied to cope with the non-convexity caused by the valve point loading effect [27]. The quasi-oppositional-based political optimizer has been implemented by [28], which has minimized the costs and emissions as well. However, with an increase in the number of units, the proposed technique requires a higher number of iterations to reach a desirable solution. The kernel search optimization algorithm and PSO with Cauchy perturbation [29] are also used to solve economic dispatch problem with non-convexity of valve point loading effect. Results indicate that the proposed methodology is robust and flexible. [30] introduced an improved mayfly algorithm incorporating levy flight to resolve the combined economic emission dispatch problem encountered in microgrids with thermal, solar, and wind generation. Although this reference, successfully minimizes total cost and emission for four different scenarios, the speed of the solution has not been discussed. Quasi-oppositional learning-based chaos-assisted sine cosine algorithm has been applied in [31] for dynamic economic emission dispatch in hybrid energy systems. In this reference, the participation of wind and solar energy resources led to a reduction of 20.84% and 8% in generation costs, respectively. [32] evaluates the implementation of the grasshopper optimization algorithm with the binary approach to solve heat and power economic emission dispatch considering the valve point loading effect, ramp-rate constraints, prohibited zones, and transmission losses. This reference has additionally introduced a new case study to examine the effectiveness of the proposed strategy. To increase convergence performance and solution speed in solving the problem of hybrid dynamic economics emissions dispatch, an improved COOT optimization algorithm is utilized [24]. Adaptive chaotic class topper optimization algorithm [33], exchange market algorithm (stock ticker EMA) [34], and chemical reaction optimization (CRO) [25] are also applied to solve combined economic emission dispatch problem.

Solving the EELD problem along with uncertainties about renewable energy sources has been associated with many challenges in recent years. Even though the literature is enriched with the optimization techniques employed to solve it but still a lot of papers are being published incorporating the problem through the variants of metaheuristic optimization techniques.

In this paper, the EELD problem has been solved by presenting an improved dynamic differential annealed optimization algorithm, which is an improved version of the DDAO algorithm. The main contributions of this work are as follows:

- An improved dynamic differential annealed optimization algorithm is proposed to enhance the performance of the conventional DDAO optimizer.
- The efficacy of the proposed algorithm to solve the EELD problem in 6-unit and 40-unit test systems is validated.
- The competence and effectiveness of the proposed IDDAO in terms of robustness, quality of the solution, and computational efficiency are compared with various methods suggested in the literature.

The remainder of this article is organized as follows:

The EELD problem is formulated in section 2. In part 3, the proposed IDDAO algorithm is introduced, followed by section 4 which discusses and evaluates the performance and effectiveness of the proposed IDDAO algorithm in 6-unit and 40-unit test systems to solve EELD with the presence of wind generation. Section 5 is dedicated to the conclusion.

II.The problem of economic emission load dispatch

With the increase in the demand for electric power generation and the reduction of power resources and the increase in pollution, the optimization issues in modern power systems have received more attention from researchers. Considering the output power of wind units, one of the most important optimization issues in power systems related to loading management on the demand side for calculation of the optimal output of generators on the generation side. In terms of the generated power of thermal units, the power plants do not have the same specifications and have different fuel costs, and are located at different distances from the load centers, and the lines that connect them to the loads have different specifications. In addition to these problems, the generation capacity of power plants is more than the total demand of loads and system losses in normal operating conditions, leading to

the issue that there are different choices for the amount of power generated by each power plant [3].

In power systems analysis, economic load dispatch has an important role, as it requires the generation units to meet the load demand while the total cost of power generation is minimized, and various non-linear and complex constraints of this process are met. In the simplest possible case, the economic load dispatch problem has the objective function of relation (1) [25].

$$F_c = \sum_{i=1}^{N} \left(a_i + b_i P_i + c_i P_i^2\right) \tag{1}$$

where the total production power of thermal power plants is represented by F_c, N represents the number of thermal units, and a_i, b_i, and c_i are the cost coefficients of each unit. Considering the effect of the steam valve, the cost function changes as follows [25].

$$\mathrm{F_c} = \sum_{\mathrm{i=1}}^{\mathrm{N}} \left(\mathrm{a_i} + \mathrm{b_i P_i} + \mathrm{c_i P_i^2} + \left|\mathrm{e_i} \times \sin\left(\mathrm{f_i}\left(\mathrm{P_i^{min}} - \mathrm{P_i}\right)\right)\right|\right) \tag{2}$$

where the coefficients f_i and e_i are related to the effect of the steam valve and the lower limit of the generation of each unit. For each thermal unit, the cost of the total emission is expressed as the following relationship [25]:

$$\boldsymbol{E_c = \sum_{i=1}^{N} 10^{-2}\left(\alpha_i + \beta_i P_i + \gamma_i P_i^2\right) + \eta_i \exp(\delta_i P_i)} \tag{3}$$

where α_i, β_i, γ_i, η_i, and δ_i are the emission coefficients of each generator. The cost of power generated by wind units in the following relationship consists of three parts [35]. The first part is related to the cost paid to the owner of wind turbines. The second part is related to the conditions in which the output power of the wind unit is greater than its estimated value, so some of the generated power will be wasted or the power of other units must be reduced by redistributing load so that the difference between the planned power of the wind unit and the output of the wind unit should be applied in the cost function. The third part of the cost function is related to the conditions in which the output power of the wind unit is less than the planned value for it. In this case, to balance the generation and consumption power, the revolving reserve capacity should be used [35].

$$F_W = \sum_{j=1}^{M} d_j w_j + k_{P,j} \int_{w_j}^{w_{r,j}} \left(w - w_j\right) f_w(w)\, dw + k_{R,j} \int_{0}^{w_j} \left(w_j - w\right) f_w(w)\, dw \tag{4}$$

In the above relation, the number of wind units and the factor of the planned power cost of the wind unit is shown by M, and d_j. In the second and third parts, $k_{P,j}$ shows the penalty factor for generating more power than the planned amount, and $k_{R,j}$ shows the penalty factor for generating less than the planned amount. In this relationship, $f_W(w)$ represents the probability density function of the output power of the wind turbine, which is calculated from the following relationships [16, 35].

$$f_W(w) = \frac{k l v_i}{c}\left(\frac{(1+\rho l)v_i}{c}\right)^{k-1} * \exp\left(-\left(\frac{(1+\rho l)v_i}{c}\right)^{k}\right) \tag{5}$$

$$l = \left(\frac{(v_r - v_i)}{v_i}\right) \tag{6}$$

$$\rho = \frac{w}{w_r} \tag{7}$$

where w is the output power of the wind turbine, v is the wind speed, ρ is the ratio of the output power to the nominal wind power, and l is the ratio of the linear range of the wind speed to the connection speed of the wind turbine.

A. Objective function

The objective function implemented in this paper is as follows:

$$T_c = Minimaze \quad (F_c + E_c + F_W) \tag{8}$$

B. System limitations

The EELD problem has several constraints. The first constraint is the balance of the power system's generation and consumption power so that the set of power generation at any time is equal to the set of consumption power and network losses. This issue is defined as the following relationship. It should be noted that the system losses are considered a function of the power generation, whose value is calculated from the following equations.

$$\sum_{i=1}^{N} P_i = P_{\mathrm{Demand}} + P_{\mathrm{Loss}} \tag{9}$$

$$P_{\mathrm{Loss}} = \sum_{i=1}^{N}\sum_{j=1}^{N} P_i B_{ij} P_j + \sum_{i=1}^{N} B_{i0} P_i + B_{00} \tag{10}$$

where, the required power is indicated by P_{Demand}, and the losses of the power system are indicated by P_{Loss}. B_{00}, B_{i0}, and B_{ij} are the network loss coefficients, and constants during normal operating conditions [28].

The second constraints indicate the minimum and maximum of generated power by thermal and wind power plants as follows [35]:

$$P_i^{min} \le P_i \le P_i^{max} 0 \le w_j \le w_{r,j} \tag{11}$$

where the upper and lower limit of the power generation of each thermal unit is indicated by P_i^{max} and P_i^{min}, and the rated power and output power for each wind unit are indicated by $w_{r,j}$, and w_j.

III.The proposed optimization algorithm

Introducing the DDAO algorithm, the proposed improved algorithm has been discussed in this section.

A. Dynamic differential annealed optimization algorithm (DDAO)

The dynamic differential annealing optimization (DDAO) algorithm presented in reference [36] is inspired by the optimal production process of two-phase annealed steel. This algorithm is effective in solving optimization problems. The processes of this algorithm are briefly described below:

The first stage includes the production of the initial population. At this stage, the initial population of solutions is generated randomly. This population includes candidate answers and general answers. In the second step, for the production of two-phase steel, the temperature decreases randomly. The random process of temperature reduction causes the formation of different phases of steel that have different energies. These energies are proportional to the value of the cost function. In the third stage, the two-phase steel cooling process is done in three ways, which are: air cooling, accelerated cooling system, and slow cooling system. This procedure can be represented by the following equation.

$$N_S^J = R_{GS} + [R_{CS}^K + R_{CS}^I] \qquad J = 1, \dots, N \qquad (12)$$

where N_S^J and R_{GS} represent the new solution in the Jth iteration and the random solution of production, respectively. R_{CS}^K and R_{CS}^I represent the randomly selected solution according to *K* and *I* index.

In the fourth stage which is related to reducing the differential temperature, the steel is exposed to the forging process. Because during the forging process, the dynamic characteristic of the hammer fluctuates between 1 and R, this process can be formulated according to the following equation.

$$F_G = \begin{cases} 1; if\ Re(I,2) = 1 \\ R[0,1]; if\ Re(I,2) = 0 \end{cases} \qquad (13)$$

where, F_G, *R*, and *Re* represent the forging parameter, the random value, and the remainder divided by 2, respectively. Forging processes and differential cooling processes are integrated in the following equation.

$$N_S^J = R_{GS} + [R_{CS}^K + R_{CS}^I] * F_G \qquad (14)$$

In the annealing process, the probability of forming a new phase with higher quality is higher than that of a weaker base. This process is defined by the acceptance probability, *Pr*, according to the following relations:

$$Pr = e^{\frac{-\Delta D}{t}} \qquad (15)$$

$$\Delta D = \frac{C(N_S^J) - C(P_M)}{C(P_M)} \qquad (16)$$

In this regard, ΔD represents the difference between the merit value of the potential responses resulting from equation (14) and the merit value corresponding to the P_M index. The values of *p* and *t* represent the population size and temperature variable, respectively. The best solution is selected by satisfying the following equation:

$$Pr > R \in [0,1] \qquad (17)$$

Based on equation (12),

$$Pr = \begin{cases} 1 & if\ t\ is\ high \\ 0 & if\ t\ is\ low \end{cases} \qquad (18)$$

The mentioned stages are repeated until reaching the stopping condition of the algorithm, which can be the maximum number of iterations or reaching a certain cost function.

B. Improved dynamic differential annealed optimization algorithm (IDDAO)

Investigations imply that the forging parameter, F_G, has a significant effect on the overall efficiency of the DDAO algorithm in solving mathematical optimization problems. For some cases, setting F_G=1 in equation (14), the random value improves while for some sets of problems, the random value becomes worse. This may reversely happen in other cases. equation (13) provides a solution for this algorithm in which F_G = 1 for half of the iterations while for other iterations F_G is a random number in the interval [0, 1]. Investigations also show that the randomness of the value of F_G causes the algorithm not to desirably perform in reaching the optimal solution; therefore, equations (13) and (14) are considered as follows.

$${F_G}^{J+1} = W_e.{F_G}^J + C.rand(N_S^{best} - N_S^J) \qquad (19)$$

$$N_S^{J+1} = N_S^J + {F_G}^{J+1} \qquad (20)$$

In these equations, W_e is defined as a weighting factor, C is a learning factor (a number between 1.5-2) and rand is a random number between [0, 1]. Also, to prevent the divergence of F_G, its final value is limited.

Equation (19) contains two terms. The first term is a ratio of the forging parameter for each phase, and its role is similar to the momentum in the neural network. The second term is proportional to the difference between the obtained answer compared to the best existing answer (N_S^{best}), which increases the speed of guiding the answer toward the optimal answer. In the proposed improved algorithm, the degree of accuracy and convergence is highly dependent on W_e, so its value is considered dynamically and linearly according to equation (21). In the beginning, by choosing large values for W_e, it is possible to find good solutions in the early stages, while in the final stages, the smallness of W_e causes better convergence.

$$w_e^{J+1} = w_e^J + [w_e^{min} + w_e^{max}]/N_{iteration} \qquad (21)$$

In this equation, w_e^{max}, w_e^{min} and $N_{iteration}$ are the maximum weighting factor, minimum weighting factor, and the number of iterations in the algorithm, respectively. The pseudo-code of this algorithm is shown in Figure 1.

IV.The EELD Based on IDDAO Algorithm

To evaluate the performance of the proposed improved IDDAO algorithm, the EELD problem for 6-unit and 40-unit standard IEEE test systems has been applied, and the results have been compared with those of PSO, GSA, and DDAO algorithms. In the proposed algorithm, the position vector is considered as the following equation.

$$x = \left[P_1, P_2, \dots, P_{N_G}\right]^T \quad (22)$$

where x is the power generated by each power plant.

In this study simulations have been done in MATLAB R2021b software, using a Core i7, 2.30 GHz processor, and a 4G Ram. The number of populations is 150, the number of iterations is 100, and the number of independent runs is 30. The best solution of all independent runs is considered the optimal solution of the algorithms.

A. *EELD for 6-unit IEEE test system*

For the standard IEEE 6-unit test system with a load demand of 2,834 p.u., cost function coefficients, the generation constraints units, thermal unit emission cost coefficients, and wind turbine data are given in Tables 1 to 3 respectively [16,37].

Considering the permitted range of power generation of power plants, the fuel cost of thermal power plants, the cost function of pollution, and two wind farms, the output power of power plants along with the optimal values of the objective function for the proposed IDDAO algorithm and PSO, GSA, DDAO algorithms are given in Table 4.

***Pseudocode*: IDDAO Algorithm**

Input: size of the population, variable temperature t, dimensions

1: Population Initialization x_K $(K = 1, \dots . N)$;
2: Initialization of cooling rate and temperature parameters;
3: Evaluating fitness for every solution;
4: Best solution $(N_S^{best}) = x_B$;
5: **While** $T < I_{MAX}$
6: Initialization of sub population R_{CS};
7: Evaluating the cost function for every R_{CS};
8: Sorting of R_{CS}
9: Select the best solution in sub population;
10: Selection of two random solutions;
11: Evaluation of N_S^J from equation (20);
12: Population sorting;
13: **for** (each solution within the population)
14: **if** (there occurs enhancement)
15: $x_K = N_S^J$
16: **else**
17: Replacement of worst solution using equation (15) and (16);
18: **end if**
19: **end for**
20: Updating of x_B ;
21: $T = T^* \, rate \, of \, cooling$
22: $T = T + 1$
23: **end while**
24: **Return** x_B

Output: Best solution

Fig. 1. The pseudo-code of the IDDAO algorithm.

TABLE 1
COST FUNCTION COEFFICIENTS OF THE 6-UNIT SYSTEM.

Unit i	Generation limits		Fuel cost coefficients				
	$P_{i,\min}$	$P_{i,\max}$	a_i	b_i	c_i	e_i	f_i
1	0.05	0.5	10	200	100	–	–
2	0.05	0.6	10	150	120	–	–
3	0.05	1.0	20	180	40	–	–
4	0.05	1.2	10	100	60	–	–
5	0.05	1.0	20	180	40	–	–
6	0.05	0.6	10	150	100	–	–

TABLE 2
EMISSION COEFFICIENTS OF THE 6-UNIT SYSTEM.

Unit i	Emission coefficients				
	α_i	β_i	γ_i	η_i	δ_i
1	4.091	−5.554	6.490	2.0e−4	2.857
2	2.543	−6.047	5.638	5.0e−4	3.333
3	4.258	−5.094	4.586	1.0e−6	8.000
4	5.326	−3.550	3.380	2.0e−3	2.000
5	4.258	−5.094	4.586	1.0e−6	8.000
6	6.131	−5.555	5.151	1.0e−5	6.667

TABLE 3
THE PARAMETERS OF WIND TURBINES.

	k	C	$c_{m,j}$	d	w_{min} (MW)	w_{max} (MW)	v_o (m/s)	v_i (m/s)	v_r (m/s)
WT1	2.2	15	100	120	10	100	45	5	15
WT2	2.2	15	100	150	10	100	45	5	15

TABLE 4
BEST SOLUTION FOR 40-UNIT TEST SYSTEM.

Outputs (MW)	PSO [14]	GSA [14]	DDAO	IDDAO
TU 1	0.02	0.1445	0.2896	0.2941
TU 2	0.3006	0.5346	0.3258	0.3325
TU 3	0.2824	0.2387	0.3501	0.3478
TU 4	0.3129	0.06	0.3625	0.3581
TU 5	0.4032	0.2548	0.3521	0.3795
TU 6	0.4161	0.2267	0.6258	0.6295
WT 1	0.8	0.5771	0.5289	0.4919
WT 2	0.302	0.8	0.2896	0.2941
Fuel cost ($/h)	624.97	588.46	571.99	565.89
EC (ton/h)	0.2043	0.2133	0.2049	0.2019
Total cost ($/h)	1836	1853	1776	1674

B. *EELD for 40-unit IEEE test system*

For the IEEE 40-unit test system, the cost function coefficients, power generation limits, and pollution coefficients are stated in Tables 5 and 6, respectively [37].

Considering the permitted range of power generation, the cost of fuel consumed by thermal power plants, and the cost function of pollution, the simulation of EELD based on the PSO, GSA, DDAO, and IDDAO algorithm was carried out, and results are summarized in Table 7.

V.Numerical results and discussions

Based on Tables 4 and 7, the total cost reduction achieved by the IDDAO algorithm compared to DDAO, GSA, and PSO algorithms for the 6-unit test system is 102 $/h, 179 $/h, and 162 $/h and for a 40-unit test system equals 1332 $/h, 7828 $/h, and 2810 $/h, respectively. Results demonstrate the superiority of the proposed algorithm in reaching the optimal solution over other algorithms.

TABLE 5
COST FUNCTION COEFFICIENTS OF THE 40-UNIT SYSTEM.

Unit i	Generation limits		Fuel cost coefficients				
	$P_{i,min}$	$P_{i,max}$	a_i	b_i	c_i	e_i	f_i
1	36	114	94.705	6.73	0.00690	100	0.084
2	36	114	94.705	6.73	0.00690	100	0.084
3	60	120	309.54	7.07	0.02028	100	0.084
4	80	190	369.03	8.18	0.00942	150	0.063
5	47	97	148.89	5.35	0.01140	120	0.077
6	68	140	222.33	8.05	0.01142	100	0.084
7	110	300	287.71	8.03	0.00357	200	0.042
8	135	300	391.98	6.99	0.00492	200	0.042
9	135	300	455.76	6.60	0.00573	200	0.042
10	130	300	722.82	12.90	0.00605	200	0.042
11	94	375	635.20	12.90	0.00515	200	0.042
12	94	375	654.69	12.80	0.00569	200	0.042
13	125	500	913.40	12.50	0.00421	300	0.035
14	125	500	1760.40	8.84	0.00752	300	0.035
15	125	500	1760.40	8.84	0.00752	300	0.035
16	125	500	1760.40	8.84	0.00752	300	0.035
17	220	500	647.85	7.97	0.00313	300	0.035
18	220	500	649.69	7.95	0.00313	300	0.035
19	242	550	647.83	7.97	0.00313	300	0.035
20	242	550	647.81	7.97	0.00313	300	0.035
21	254	550	785.96	6.63	0.00298	300	0.035
22	254	550	785.96	6.63	0.00298	300	0.035
23	254	550	794.53	6.66	0.00284	300	0.035
24	254	550	794.53	6.66	0.00284	300	0.035
25	254	550	801.32	7.10	0.00277	300	0.035
26	254	550	801.32	7.10	0.00277	300	0.035
27	10	150	1055.10	3.33	0.52124	120	0.077
28	10	150	1055.10	3.33	0.52124	120	0.077
29	10	150	1055.10	3.33	0.52124	120	0.077
30	47	97	148.89	5.35	0.01140	120	0.077
31	60	190	222.92	6.43	0.00160	150	0.063
32	60	190	222.92	6.43	0.00160	150	0.063
33	60	190	222.92	6.43	0.00160	150	0.063
34	90	200	107.87	8.95	0.00010	200	0.042
35	90	200	116.58	8.62	0.00010	200	0.042
36	90	200	116.58	8.62	0.00010	200	0.042
37	25	110	307.45	5.88	0.01610	80	0.098
38	25	110	307.45	5.88	0.01610	80	0.098
39	25	110	307.45	5.88	0.01610	80	0.098
40	242	550	647.83	7.97	0.00313	300	0.035

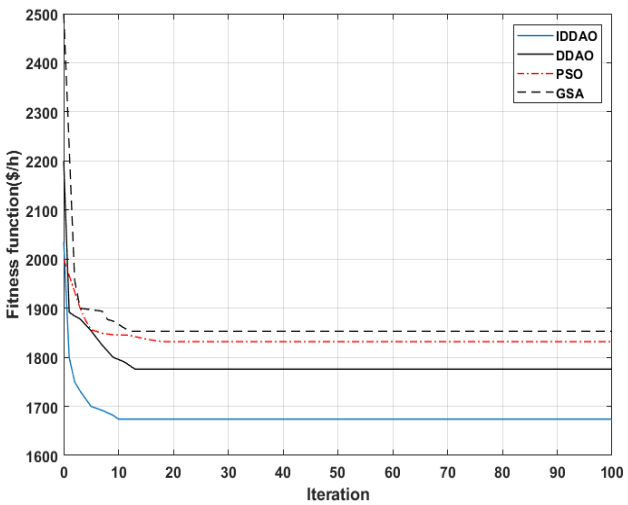


Fig 2. Convergence behaviors of fitness function for the 6-unit test system.

The convergence curve of the proposed improved algorithm for the best solution obtained for the 6-unit and 40-unit test systems is shown in Figures 2 and 3. The IDDAO algorithm has converged to a more optimal solution after a smaller number of iterations compared to other algorithms. PSO, GSA, and DDAO algorithms have a lower convergence speed and are stuck in the local solution. Hence, it can be concluded that the performance of the proposed algorithm is satisfactory for solving the EELD problem.

TABLE 6
EMISSION COEFFICIENTS OF THE 40-UNIT SYSTEM.

Unit i	Emission coefficients				
	α_i	β_i	γ_i	η_i	δ_i
1	60	−2.22	0.0480	1.3100	0.05690
2	60	−2.22	0.0480	1.3100	0.05690
3	100	−2.36	0.0762	1.3100	0.05690
4	120	−3.14	0.0540	0.9142	0.04540
5	50	−1.89	0.0850	0.9936	0.04060
6	80	−3.08	0.0854	1.3100	0.05690
7	100	−3.06	0.0242	0.6550	0.02846
8	130	−2.32	0.0310	0.6550	0.02846
9	150	−2.11	0.0335	0.6550	0.02846
10	280	−4.34	0.4250	0.6550	0.02846
11	220	−4.34	0.0322	0.6550	0.02846
12	225	−4.28	0.0338	0.6550	0.02846
13	300	−4.18	0.0296	0.5035	0.02075
14	520	−3.34	0.0512	0.5035	0.02075
15	510	−3.55	0.0496	0.5035	0.02075
16	510	−3.55	0.0496	0.5035	0.02075
17	220	−2.68	0.0151	0.5035	0.02075
18	222	−2.66	0.0151	0.5035	0.02075
19	220	−2.68	0.0151	0.5035	0.02075
20	220	−2.68	0.0151	0.5035	0.02075
21	290	−2.22	0.0145	0.5035	0.02075
22	285	−2.22	0.0145	0.5035	0.02075
23	295	−2.26	0.0138	0.5035	0.02075
24	295	−2.26	0.0138	0.5035	0.02075
25	310	−2.42	0.0132	0.5035	0.02075
26	310	−2.42	0.0132	0.5035	0.02075
27	360	−1.11	1.8420	0.9936	0.04060
28	360	−1.11	1.8420	0.9936	0.04060
29	360	−1.11	1.8420	0.9936	0.04060
30	50	−1.89	0.0850	0.9936	0.04060
31	80	−2.08	0.0121	0.9142	0.04540
32	80	−2.08	0.0121	0.9142	0.04540
33	80	−2.08	0.0121	0.9142	0.04540
34	65	−3.48	0.0012	0.6550	0.02846
35	70	−3.24	0.0012	0.6550	0.02846
36	70	−3.24	0.0012	0.6550	0.02846
37	100	−1.98	0.0950	1.4200	0.06770
38	100	−1.98	0.0950	1.4200	0.06770
39	100	−1.98	0.0950	1.4200	0.06770
40	220	−2.68	0.0151	0.5035	0.02075

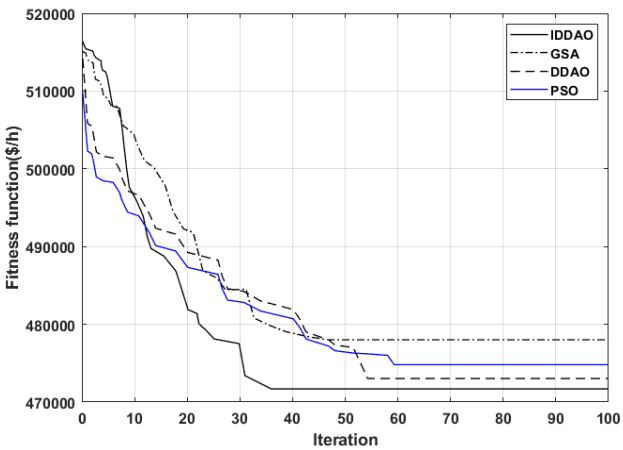


Fig 3. Convergence behaviors of fitness function for the 40-unit test system.

To check the compatibility of the proposed improved algorithm, the values of the best solution, the worst solution, and the average solution along with the standard deviation (STD) considering 30 independent runs for PSO, GSA, DDAO, and IDDAO algorithms for 6-unit and 40-unit test system are given in Table 8. Results show that the standard deviation of the optimal solutions obtained by the IDDAO algorithm is smaller than other algorithms, which indicates its higher robustness and stability to different parameters of algorithms. Comparing the average time required to run the proposed algorithm with other algorithms, it can also be concluded that the solution speed of the IDDAO algorithm surpasses those of other algorithms.

TABLE 7
BEST SOLUTION FOR THE 40-UNIT TEST SYSTEM.

Outputs (MW)	PSO [25]	GSA [25]	DDAO	IDDAO
TU 1	114.0000	112.0124	113.2849	112.152051
TU 2	106.2356	102.0546	112.9873	110.6585
TU 3	118.2564	65.3624	118.9851	119.1124
TU 4	182.6634	189.0000	172.3033	172.5485
TU 5	97.0000	81.4570	96.9899	97.0000
TU 6	102.1987	140.0000	120.7941	118.5487
TU 7	134.0531	178.9692	297.9866	240.5487
TU 8	292.2256	261.0062	298.0531	299.1473
TU 9	290.1165	144.0886	289.9833	290.2548
TU 10	133.8941	268.6632	141.9755	138.2546
TU 11	101.2368	262.5468	290.8744	278.2351
TU 12	154.7289	338.0018	295.8873	299.2548
TU 13	308.2166	357.2549	428.9894	425.365
TU 14	368.3325	402.6638	421.9738	425.9854
TU 15	371.2215	481.0378	430.9884	424.2154
TU 16	381.2681	426.7826	416.9725	398.8547
TU 17	416.2678	468.2594	449.7641	448.6354
TU 18	482.9655	411.0338	451.0875	451.5241
TU 19	506.2201	452.7650	463.4483	466.9878
TU 20	516.2468	485.3309	424.9509	420.6662
TU 21	502.7719	410.7262	418.7534	417.2154
TU 22	435.4741	446.8291	438.5739	448.754
TU 23	482.1264	410.5897	415.5968	442.1187
TU 24	538.0536	429.6281	429.0083	430.2522
TU 25	545.0089	355.4008	442.8693	447.6584
TU 26	426.2278	497.2740	421.7939	400.2215
TU 27	121.2238	110.0368	29.6631	38.2654
TU 28	126.0343	103.0191	20.9746	50.2654
TU 29	106.2268	135.4494	28.2672	102.5245
TU 30	96.0892	67.2264	95.0546	90.2154
TU 31	172.0086	160.2286	183.1982	178.2654
TU 32	181.6147	156.2273	169.8629	168.1145
TU 33	170.1243	162.1683	183.2297	175.2654
TU 34	191.3843	150.6824	184.8734	190.2487
TU 35	189.5371	163.4874	190.2495	192.3659
TU 36	172.0034	173.7898	197.8624	186.2658
TU 37	97.6206	99.3548	102.6921	99.1212
TU 38	94.4833	96.1757	103.2485	104.3256
TU 39	86.0853	99.0471	100.2394	95.3298
TU 40	462.2264	484.1232	432.0381	420.6965
WT 1	46.0468	66.2162	41.5432	45.2185
WT 2	80.0468	94.0293	32.1267	40.2354
Fuel cost($/h)	142 068	143562.68	141241.46	141025.41
EC (ton/h)	178432	180028.22	171953.76	171902.52
Total cost($/h)	474934	479405	473456	472124

TABLE 8
STATISTICAL COMPARISON BETWEEN IDDAO AND DIFFERENT.

Test systems	Algorithms	Best($/h)	Worst($/h)	Average($/h)	STD	Average of run time(s)
6-unit	PSO	1836	1880	1856	15.1471	3.542
	GSA	1853	1914	1886	20.5487	4.754
	DDAO	1776	1831	1793	11.2514	4.124
	IDDAO	1674	1724	1685	7.2135	3.562
40-unit	PSO	474934	495356	481108	6451.7854	26.2541
	GSA	479405	503375	489951	8721.2145	34.524
	DDAO	473456	492394	479610	1425.5215	34.854
	IDDAO	472124	490064	478733	501.2341	25.43

VI.Conclusions

In this article, to solve the EELD problem with the participation of wind farms, a novel strategy based on the IDDAO algorithm has been introduced. The limitations such as generation-demand balance, uncertainty, and the non-convexity of cost functions for thermal power plants have also been considered. The performance of the proposed algorithm to solve the EELD problem for two standard test systems of 6-unit and 40-unit were compared and evaluated. The minimum reduction of the total cost function obtained by the proposed IDDAO algorithm Compared to the DDAO, GSA, and PSO algorithms was 102 $/h and 1332 $/h for 6-unit and 40-unit test systems, respectively. These results demonstrate that the new approach developed in this paper can efficiently reduce the expected cost of the integrated thermal-wind system. Moreover, to check the robustness and stability of the proposed algorithm, the EELD problem was implemented for 30 independent runs. The results indicate the superiority of the proposed algorithm in the speed of convergence and proper performance compared to other studied algorithms.

REFERENCES

[1] S. Abbas, G. Ali, F. M. Qamer, and S. M. Irteza, "Associations of air pollution concentrations and energy production dynamics in Pakistan during lockdown," Environmental Science and Pollution Research, vol. 29, no. 23, pp. 35036-35047, 2022.

[2] A. B. Kunya, A. S. Abubakar, and S. S. Yusuf, "Review of economic dispatch in multi-area power system: State-of-the-art and future prospective," Electric Power Systems Research, vol. 217, p. 109089, 2023.

[3] J. Zhu, Optimization of power system operation. John Wiley & Sons, 2015.

[4] H. Liang, Y. Liu, Y. Shen, F. Li, and Y. Man, "A hybrid bat algorithm for economic dispatch with random wind power," IEEE Transactions on Power Systems, vol. 33, no. 5, pp. 5052-5061, 2018.

[5] M. Kheshti, L. Ding, S. Ma, and B. Zhao, "Double weighted particle swarm optimization to non-convex wind penetrated emission/economic dispatch and multiple fuel option systems," Renewable Energy, vol. 125, pp. 1021-1037, 2018.

[6] Z. Lin, H. Chen, Q. Wu, W. Li, M. Li, and T. Ji, "Mean-tracking model based stochastic economic dispatch for power systems with high penetration of wind power," Energy, vol. 193, p. 116826, 2020.

[7] K. Alawode, A. Jubril, L. Kehinde, and P. O. Ogunbona, "Semidefinite programming solution of economic dispatch problem with non-smooth, non-convex cost functions," Electric Power Systems Research, vol. 164, pp. 178-187, 2018.

[8] B. Qiao and J. Liu, "Multi-objective dynamic economic emission dispatch based on electric vehicles and wind power integrated system using differential evolution algorithm," Renewable Energy, vol. 154, pp. 316-336, 2020.

[9] S. Hazra and P. K. Roy, "Quasi-oppositional chemical reaction optimization for combined economic emission dispatch in power system considering wind power uncertainties," Renewable energy focus, vol. 31, pp. 45-62, 2019.

[10] Y. Zhang, K. Liu, X. Liao, L. Qin, and X. An, "Stochastic dynamic economic emission dispatch with unit commitment problem considering wind power integration," International Transactions on Electrical Energy Systems, vol. 28, no. 1, p. e2472, 2018.

[11] W. Sheng, R. Li, T. Yan, M.-L. Tseng, J. Lou, and L. Li, "A hybrid dynamic economics emissions dispatch model: Distributed renewable power systems based on improved COOT optimization algorithm," Renewable Energy, 2023.

[12] H. Lotfi, "A Multiobjective Evolutionary Approach for Solving the Multi-Area Dynamic Economic Emission Dispatch Problem Considering Reliability Concerns,"

Sustainability, vol. 15, no. 1, p. 442, 2022.
[13] K. D. Bodha, V. K. Yadav, and V. Mukherjee, "Combined Economic Emission Dispatch of a Hybrid Energy System Using Quantum-Inspired Tidal Firefly Algorithm," in Smart Energy and Advancement in Power Technologies: Select Proceedings of ICSEAPT 2021 Volume 1: Springer, 2022, pp. 313-324.
[14] S. Jiang, Z. Ji, and Y. Wang, "A novel gravitational acceleration enhanced particle swarm optimization algorithm for wind–thermal economic emission dispatch problem considering wind power availability," International Journal of Electrical Power & Energy Systems, vol. 73, pp. 1035-1050, 2015.
[15] C. Shilaja, G. Nalinashini, N. Balaji, and K. Sujatha, "Combined Economic-Emission Dispatch of External Optimization Integrating Renewable Energy Resources Stochastic Wind and Solar," ECS Transactions, vol. 107, no. 1, p. 7325, 2022.
[16] M. T. Hagh, S. M. S. Kalajahi, and N. Ghorbani, "Solution to economic emission dispatch problem including wind farms using Exchange Market Algorithm Method," Applied Soft Computing, vol. 88, p. 106044, 2020.
[17] M. A. El-ziady, E. Elbeltagi, I. El-kalla, and M. Abouhawwash, "Solving the economic load dispatch based on NSGA-II and RNSGA-II," International Journal of Renewable Energy Research (IJRER), vol. 12, no. 1, pp. 458-465, 2022.
[18] O. Oyajide Depo, A. Olanite Olanrewaju, and E. Omotayo Mayowa, "Development of hybridized evolutionary algorithm for solving economic environmental dispatch problem in Nigerian power system," 2022.
[19] A. Xia, X. Wu, and Y. Bai, "A new multi - objective hybrid optimization algorithm for wind - thermal dynamic economic emission power dispatch," International Transactions on Electrical Energy Systems, vol. 31, no. 8, p. e12966, 2021.
[20] A. Xia and X. Wu, "A hybrid multi-objective optimization algorithm for economic emission dispatch considering wind power uncertainty," Iranian Journal of Science and Technology, Transactions of Electrical Engineering, vol. 45, pp. 1277-1293, 2021.
[21] L. R. Robert and L. R. S. Ravi Singh, "Economic emission dispatch of hydro - thermal - wind using CMQLSPSN technique," IET Renewable Power Generation, vol. 14, no. 14, pp. 2680-2692, 2020.
[22] T. Guesmi, A. Farah, I. Marouani, B. Alshammari, and H. H. Abdallah, "Chaotic sine–cosine algorithm for chance - constrained economic emission dispatch problem including wind energy," IET Renewable Power Generation, vol. 14, no. 10, pp. 1808-1821, 2020.
[23] M. Karimzadeh Parizi, F. Keynia, and A. Khatibi Bardsiri, "Woodpecker mating algorithm for optimal economic load dispatch in a power system with conventional generators," International Journal of Industrial Electronics Control and Optimization, vol. 4, no. 2, pp. 221-234, 2021.
[24] W. Sheng, R. Li, T. Yan, M.-L. Tseng, J. Lou, and L. Li, "A hybrid dynamic economics emissions dispatch model: Distributed renewable power systems based on improved COOT optimization algorithm," Renewable Energy, 2023.
[25] S. Hazra and P. K. Roy, "Quasi-oppositional chemical reaction optimization for combined economic emission dispatch in power system considering wind power uncertainties," Renewable energy focus, vol. 31, pp. 45-62, 2019.
[26] H. Sharifzadeh, "Solving economic load dispatch by a new hybrid optimization method," International Journal of Industrial Electronics Control and Optimization, vol. 3, no. 4, pp. 469-474, 2020.
[27] M. Azeem et al., "Combined Economic Emission Dispatch in Presence of Renewable Energy Resources Using CISSA in a Smart Grid Environment," Electronics, vol. 12, no. 3, p. 715, 2023.
[28] V. Basetti et al., "Economic emission load dispatch problem with valve-point loading using a novel quasi-oppositional-based political optimizer," Electronics, vol. 10, no. 21, p. 2596, 2021.
[29] R. Dong, L. Ma, H. Chen, A. A. Heidari, and G. Liang, "Hybrid kernel search and particle swarm optimization with Cauchy perturbation for economic emission load dispatch with valve point effect," Frontiers in Energy Research, vol. 10, p. 1061408, 2023.
[30] K. Nagarajan, A. Rajagopalan, S. Angalaeswari, L. Natrayan, and W. D. Mammo, "Combined economic emission dispatch of microgrid with the incorporation of renewable energy sources using improved mayfly optimization algorithm," Computational Intelligence and Neuroscience, vol. 2022, 2022.
[31] K. Dasgupta, P. Kumar Roy, and V. Mukherjee, "A Novel Quasi-Oppositional Learning-Based Chaos-Assisted Sine Cosine Algorithm for Hybrid Energy Integrated Dynamic Economic Emission Dispatch," IETE Journal of Research, pp. 1-28, 2023.
[32] Y. Sharifian and H. Abdi, "Solving multi-zone combined heat and power economic emission dispatch problem considering wind uncertainty by applying grasshopper optimization algorithm," Sustainable Energy Technologies and Assessments, vol. 53, p. 102512, 2022.
[33] A. Srivastava and D. K. Das, "An adaptive chaotic class topper optimization technique to solve economic load dispatch and emission economic dispatch problem in power system," Soft Computing, vol. 26, no. 6, pp. 2913-2934, 2022.
[34] M. T. Hagh, S. M. S. Kalajahi, and N. Ghorbani, "Solution to economic emission dispatch problem including wind farms using Exchange Market Algorithm Method," Applied Soft Computing, vol. 88, p. 106044, 2020.
[35] J. Hetzer, C. Y. David, and K. Bhattarai, "An economic dispatch model incorporating wind power," IEEE Transactions on energy conversion, vol. 23, no. 2, pp. 603-611, 2008.
[36] H. N. Ghafil and K. Jármai, "Dynamic differential annealed optimization: New metaheuristic optimization algorithm for engineering applications," Applied Soft Computing, vol. 93, p. 106392, 2020.
[37] D. Zou, S. Li, Z. Li, and X. Kong, "A new global particle swarm optimization for the economic emission dispatch with or without transmission losses," Energy Conversion and Management, vol. 139, pp. 45-70, 2017.

Mehdi Shafiee was born in Kerman, Iran in 1983. He received the B.Sc. degree from Shahid Bahonar University of Kerman, Kerman, Iran, in 2006, the M.Sc. degree from Amirkabir University of Technology, Tehran, Iran, in 2008, and the Ph.D. degree from Ferdowsi University of Mashhad, Mashhad, Iran, in 2018, all in Power Electrical Engineering. He is currently an Assistant Professor in

the Department of Electrical Engineering, Technical and Vocational University, Kerman, Iran. His research interests are power system planning and operation, Smart grids, Renewable energies, and Power management.

Abbas-Ali Zamani was born in Isfahan, Iran in 1986. He received the B.Sc. degree in Electronic Engineering from Hakim Sabzevari University, Sabzevar, Iran 2009. He received the M.Sc. degree from Isfahan University of Technology, Isfahan, Iran 2011, and the Ph.D. degree from the University of Sistan and Baluchestan, Sistan and Baluchestan, Iran, in 2018, in Control Engineering. He is an Assistant Professor in the Department of Electrical Engineering, Technical and Vocational University, Isfahan, Iran. His research interests are Seismic Control, Power System Control, Smart grids, Renewable energies, and Artificial intelligence.

Mehdi Sajadini was born in Kerman, Iran in 1983. He received the B.Sc. degree from Shahid Bahonar University of Kerman, Kerman, Iran, in 2007 and the M.Sc. degree from Bu-Ali Sina University, Hamadn, Iran 2010, both in Electrical Engineering- Power. He is a lecturer in the Department of Electrical Engineering, Technical and Vocational University, Kerman, Iran. His research interests are Power System Control, Smart grids, Renewable energies, and the application of Artificial intelligence in modern power systems.

IECO

This page intentionally left blank.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Investigation of Deep Learning Optimization Algorithms in Scene Text Detection

Zobeir Raisi [1,2] | John Zelek [1]

Systems Design Engineering Department, University of Waterloo, Waterloo, Ontario, Canada.[1]
Chabahar Maritime University, Chabahar, Iran.[2]
Corresponding author's email: zraisi@uwaterloo.ca

Article Info

Article type:
Research Article

Article history:
Received: 16-May-2023
Received in revised form: 01-Aug-2023
Accepted: 13-Augu-2023
Published online: 25-Aug-2023

Keywords:
Deep Learning,
Loss Function,
Optimization,
Scene Text Detection,
Attention Module

ABSTRACT

Scene text detection frameworks heavily rely on optimization methods for their successful operation. Choosing an appropriate optimizer is essential to performing recent scene text detection models. However, current deep learning methods often employ various optimization algorithms and loss functions without explicitly explaining their selections. This paper presents a simple segmentation-based text detection pipeline capable of handling arbitrary-shaped text instances in wild images. We explore the effectiveness of well-known deep-learning optimizers to enhance the pipeline's capabilities. Additionally, we introduce a novel Segmentation-based Attention Module (SAM) that enables the model to capture long-range dependencies of multi-scale feature maps and focus more accurately on regions likely to contain text instances. The performance of the proposed architecture is extensively evaluated through ablation experiments, exploring the impact of different optimization algorithms and the introduced SAM block. Furthermore, we compare the final model against state-of-the-art scene text detection techniques on three public benchmark datasets, namely ICDAR15, MSRA-TD500, and Total-Text. Our experimental results demonstrate that the Focal Loss (FL) combined with the Stochastic Gradient Descent (SGD) + Momentum optimizer using a poly learning-rate policy achieves a more robust and generalized detection performance than other optimization strategies. Moreover, our utilized architecture, empowered by the proposed SAM block, significantly enhances the overall detection performance, achieving competitive H-mean detection scores while maintaining superior efficiency in terms of Frames Per Second (FPS) compared to recent techniques. Our findings shed light on the importance of selecting appropriate optimization strategies and demonstrate the effectiveness of our proposed SAM in scene text detection tasks.

I. Introduction

Text Reading from wild images is a common research topic in computer vision with various valuable applications [1-4]. It has numerous practical applications such as optical character recognition, multi-lingual translation, image search, etc. [5-6]. Reading text from images in the environment can be broken down into two parts: (1) text detection, which focuses on locating the text within the image, and (2) text recognition, which involves converting the located text or individual word image into a string. This paper focuses on text detection, which is more complex than recognition due to wild images' variety of text shapes and complex backgrounds.

Before the deep learning era, text detection techniques relied on connected components or sliding window-based techniques, which used handcrafted features such as MSER [7,8] or SWT [9]. These techniques had several significant areas for improvement: (1) they were designed only for detecting individual characters or components, making it challenging to identify regional context information and resulting in low recall performance. (2) They required multiple post-processing steps for text detection. (3) They were limited to horizontal text and could not handle multi-oriented text instances.

Recently, deep learning-based techniques are demonstrated superior performance to traditional classical machine learning-based methods for detecting various challenging text instances appearing in wild images. These deep-learning methods often utilize popular frameworks inherited from object detection communities, such as SSD [10], YOLO [11], Faster R-CNN [12], or object segmentation architectures such as FCN [13] and Mask R-CNN [14]. Deep learning-based text detectors mainly detect text at the word level, and they usually struggle with finding curved, very long, or abnormally shaped word instances by using only a single bounding box [15]. Improving the accuracy of text bounding boxes without significantly increasing the number of proposals, particularly in the case of curved or irregular-shaped text, is a crucial challenge in scene text detection.

Optimization plays a vital role in deep learning problems, and the objective is to find a way to optimize efficiency and determine the parameters that minimize the loss function [48, 49, 61]. Optimization algorithms [16-19] are defined by their update rule, which is regulated by hyper-parameters influencing its behavior (*e.g.*, the learning rate). However, no clear theory exists on which optimizer is best suited for scene text detection.

Attention, in the context of deep learning and natural language processing, refers to a mechanism that allows a model to focus on specific parts of the input data that are more relevant to the task. It enables the model to assign weights to different input parts, giving higher importance to the more informative elements. In computer vision, attention mechanisms are used successfully in tasks like image captioning, object detection, and visual question answering [50, 51]. In text detection, some recent methods [52, 53] benefited the attention mechanism in their models and achieved superior performance in benchmark datasets.

This paper utilizes a straightforward pipeline for scene text detection, which employs a ResNet [20] backbone for feature extraction followed by a Feature Pyramid Network (FPN) to obtain multi-scale feature maps. Furthermore, we design a specific attention module and insert it after FPN, namely the Segmentation-based Attention Module (SAM), to focus on different regions of the feature maps that are more likely to contain text. This framework makes the final model fast and capable of detecting arbitrary text shapes, such as curved or multi-oriented text. This paper also addresses the issue of various optimizers used in scene text detection architectures by examining the loss functions and optimization techniques commonly used in deep learning for object detection and segmentation. Moreover, it selects the best loss function and optimization method for the leveraged scene text detection framework.

Our main contributions are: (1) we comprehensively compare and analyze deep learning optimization algorithms to determine their effectiveness in text detection applications. (2) we design an attention-based module, SAM, to boost the detection performance of the utilized detector. (3) we conduct several ablation experiments to evaluate the effect of the designed block with the baseline architecture. (4) we compare the final proposed model with state-of-the-art text detection techniques on three well-known benchmark datasets.

II. Related Work

This section briefly overviews the deep learning-based techniques for scene text detection and examines the various loss functions and optimization methods used in these approaches.

C. Text Detection

Generally, scene text detection methods can be categorized as follows:

Regression-based (RB): Several recent deep learning-based scene text detection techniques [21, 22] used general object detection frameworks, like SSD [10] or Faster R-CNN [12]. These methods treat text regions as objects and predict candidate bounding boxes from them directly. For instance, TextBoxes [22] altered the SSD [10] kernels by using long default anchors and filters to handle the diverse aspect ratios of text instances and detect various text shapes. Unlike TextBoxes, the deep matching prior network (DMPNet) [23] employed sliding windows with quadrilateral shapes to catch text with multiple orientations.

Several regression-based methods [24, 25] attempted to tackle the detection challenges posed by rotated and arbitrarily shaped text. For example, EAST [21] introduced a fast and precise text detector that uses dense predictions and NMS to detect multi-oriented text in an image without needing manual anchors. Regression-based methods generally have a straightforward post-processing system for handling multi-oriented text. However, due to limitations in their structure, it is challenging to produce accurate bounding boxes for text instances with arbitrary shapes.

Segmentation-based (SB): Some techniques [26-28, 34] use pixel-level classification to detect text regions on a word or character level. These methods usually adopt a segmentation framework such as FCN [13] or Mask R-CNN [14]. For instance, Zhang et al. [26] applied FCN to generate a salience map of text regions. TextSnake [28] used FCN as a

foundation and identified text instances by detecting and piecing together local parts.

The previously mentioned methods are designed to detect text in images at the word level. However, detecting text using word instances is difficult, as individual characters may have arbitrary shapes. As a result, some recent text detection techniques shifted towards character-level detection [15, 29, 30]. For instance, [29] generates multi-oriented text bounding boxes using a saliency map of text regions obtained from a segmentation network and character-level annotations. Another example is SegLink [30], which identifies small text segments in images and combines them to form word boxes through additional post-processing. CRAFT [15] uses a weakly supervised approach to detect individual characters in the arbitrarily shaped text, leading to state-of-the-art results in benchmark datasets. These recent methods, which adopt a segmentation framework, tend to better detect multi-oriented text than regression-based methods. However, they also come with more complex and time-consuming post-processing.

Several approaches [31, 32] aimed to simplify the architecture and enhance computation speed. Inspired by [32, 33], our method offers a streamlined pipeline that can accurately detect arbitrarily shaped text more efficiently than the various state-of-the-art methods [15, 21, 27, 28, 34, 52-60].

Attention-based Detectors: The attention mechanism helps the model to focus on the relevant parts of the image that are likely to contain text, improving the localization accuracy of the text detection system. By adapting the attention module to the specific characteristics of the text detection task, the model can achieve better results than the architecture without attention.

Many recent works [50-52] utilized attention in their pipeline to detect arbitrarily shaped text and object instances and achieved superior performance in benchmarks [50, 51]. For example, Hou et al. in [50] utilized coordinate information to selectively attend to relevant spatial locations in feature maps, improving model accuracy and reducing computational complexity. Tang et al. [52] presented a transformer-based approach for scene text detection, using representative features to avoid background disturbance and reduce computational costs. Their method achieves state-of-the-art results on popular datasets by effectively grouping components corresponding to text instances, enabling easy bounding box extraction without post-processing**.**

D. Loss Function for Deep Learning

The loss function represents our model's disagreement with the actual and predicted labels during the training phase. By evaluating the extent of this disagreement, we gain valuable insights into the model's learning progress and its ability to capture the underlying patterns within the data effectively. This section examines the standard loss functions utilized in recent deep learning-based scene text detection and object detection architectures.

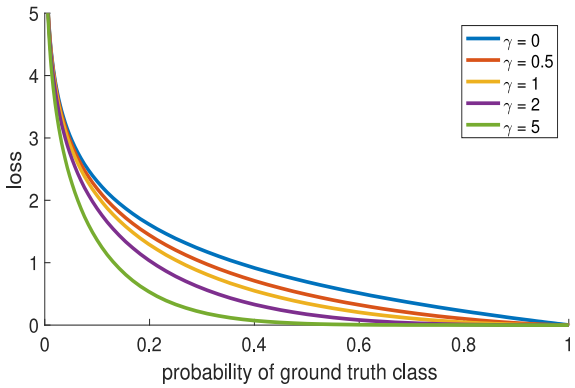


Fig. 1. Focal loss with different values of Gammas (γ).

Cross-Entropy (CE) Loss: This function evaluates the accuracy of a classification model, which gives an output between 0 and 1 in the form of probability. It is defined as:

$$\mathrm{CE}(p,y)=\begin{cases}-\log(p) & \text{if } y=1\\ -\log(1-p) & \text{otherwise.}\end{cases} \quad (1)$$

where $y \in \{\pm 1\}$ represents the ground-truth class and the $p \in [0,1]$ denotes the estimated probability of model for the class with label $y = 1$. For simplicity, we can define p_t as follows:

$$p_t=\begin{cases}p & \text{if } y=1\\ 1-p & \text{otherwise,}\end{cases} \quad (2)$$

Therefore, we can calculate $\mathrm{CE}(p,y)=\mathrm{CE}(p_t)=-\log(p_t)$. As evident from Figure 1, the cross-entropy loss has a drawback, shown in its plot (top), that it incurs a loss even for easily classified examples ($p_t > 0.5$) of non-trivial magnitude. When added over many straightforward examples, this could lead to small loss values, overpowering the rare class [35].

Focal Loss: The Focal Loss (FL), introduced in [35], is a modified version of the Cross-Entropy loss that assigns varying levels of importance to individual samples based on their classification error. The focal loss includes a modulation factor $(1-p_t)^{\gamma}$, controlled by the focusing parameter $\gamma \geq 0$, which is added to the Cross-Entropy loss. As shown in Figure 1, $\gamma = 0$ is utilized for Binay Cross Entropy (BCE) loss, $\gamma > 0$ is known as focal loss, and it outperforms BCE loss in object detection problems. The Focal Loss is defined as follows (more detail in Figure 1):

$$\mathrm{FL}(p_t)=-\alpha_t(1-p_t)^{\gamma}\log(p_t) \quad (3)$$

A balanced version of the Focal Loss is used in the implementation time with a default value of $\alpha = 0.9$. In this paper, we use $\gamma = 2$ in our experiments.

E. Optimization Algorithms for Deep Learning

The optimization algorithm minimizes or maximizes a given function, $f(x)$, in deep learning-based methods. Optimization aims to train the network by minimizing the loss function L,

representing the difference between the predicted value y' and the actual value y. The predicted value y' is obtained

TABLE 1.
PARAMETERS AND THEIR SPECIFICATIONS THAT ARE USED IN OPTIMIZATION EQUATIONS.

Parameter	Specification
t	time step
w	weight
α	learning rate
$\partial\mathcal{L}/\partial w$	gradient of $\mathcal{L}$, Loss function, *w.r.t.* to w

through forward propagation, which uses the weights (w) and biases (b) of the network. Optimization algorithms update the values of W and b to minimize the cost function L. This section reviews some commonly used optimizers. A summary of the parameters and specifications used in these algorithms is provided in Table 1.

Given a function of $y = f(x)$, an optimization algorithm helps minimize or maximize the value of $f(x)$. In deep learning-based architectures, optimizers are used to train the frameworks by minimizing the loss function $\mathcal{L}$, which is defined as follows :

$$\mathcal{L}(w,b) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}_i\,(y_i', y_i) \tag{4}$$

The value of $\mathcal{L}$ is defined as the mean loss of the predicted value y' by the detector and the actual value y of the ground- truth.

Using the network's weights w and biases b during the forward propagation step of training, the values y' are computed. Furthermore, by updating the trainable parameters, w and b, we can minimize the value of the loss function $\mathcal{L}$ using different optimization algorithms. In the rest of this section, we review commonly used optimizers used in deep learning frameworks. Table 1 shows the commonly used parameters for explaining optimization algorithms.

Gradient Descent: is a first-order optimization algorithm used to minimize the cost function $\mathcal{L}$ and find the optimal values of the weight matrix w and bias b. It updates the parameters after going through the entire training data, which can be a challenge if the data is too large to fit in memory [36]. Mathematically it can be defined as:

$$w_{t+1} = w_t - \alpha\frac{\partial\mathcal{L}}{\partial w_t} \tag{5}$$

In Stochastic Gradient Descent (SGD), instead of evaluating the loss function and gradient using all the training examples, a small subset, known as a mini-batch (B), is randomly selected and used to compute an approximation of the total sum and the actual gradient in each iteration.

$$\frac{\partial\mathcal{L}(w)}{\partial w} = \frac{1}{B}\sum_{k=1}^{B}\frac{\partial\mathcal{L}_k(w)}{\partial(w)}, \tag{6}$$

$$w_{t+1} = w_t - \alpha\frac{\partial\mathcal{L}}{\partial w_t} \tag{7}$$

Mini-Batch Gradient Descent addresses the limitations of Gradient Descent by dividing the entire training data into smaller subsets, known as mini-batches, of sizes 8, 16, 32, or 64. These mini-batches are then used to train the network in a series of iterations. SGD offers several benefits compared to traditional Gradient Descent, such as more accurate data estimation, smoother convergence, a larger learning rate, and faster training. However, SGD also has some drawbacks, including the tendency to follow shallow dimensions in rapidly changing loss functions, getting stuck in local minima or saddle points (which are more common in high-dimensional problems like scene text detection), and the noise in gradient calculations from using mini-batches [19].

SGD+Momentum: To address the limitations of Stochastic Gradient Descent (SGD), adding a term called momentum [16] can be helpful. Thus we can define it as follows:

$$\begin{aligned} v_{t+1} &= \rho v_t + \frac{\partial\mathcal{L}}{\partial w_t} \\ w_{t+1} &= w_t - \alpha v_{t+1} \end{aligned} \tag{8}$$

where ρ denotes the friction constant; typically, its default value is 0.9 or 0.99. $v(t)$ is the velocity that builds up as a running mean of gradients.

AdaGrad: Adaptive gradient or AdaGrad [17], dynamically updates the learning rate at each update and for each weight individually, in which the gradient component remains unchanged like in SGD and is defined as follows:

$$\begin{aligned} w_{t+1} &= w_t - \frac{\alpha}{\sqrt{v_t+\epsilon}}\cdot\frac{\partial\mathcal{L}}{\partial w_t}, \\ v_t &= v_{t-1} + [\frac{\partial\mathcal{L}}{\partial w_t}]^2 \end{aligned} \tag{9}$$

In the above equation, ϵ shows a small positive value. However, the main weakness of AdaGrad is its learning rate, α, which is continuously decreasing and decaying. This can lead to slow or stalled training convergence, particularly for deep neural networks and complex models.

RMSprop: An alternative method for optimizing adaptive learning rates is RMSprop [18], designed to enhance AdaGrad [17]. Unlike AdaGrad [17], which accumulates the squared gradients, RMSprop employs an exponential moving average of these gradient values. RMSprop has now become a common choice, like momentum, for updating the learning rate component in the majority of optimization algorithms.

$$\begin{aligned} w_{t+1} &= w_t - \frac{\alpha}{\sqrt{v_t+\epsilon}}\cdot\frac{\partial\mathcal{L}}{\partial w_t}, \\ v_t &= \beta v_{t-1} + (1-\beta)[\frac{\partial\mathcal{L}}{\partial w_t}]^2 \end{aligned} \tag{10}$$

In the above equation, β is a hyperparameter with a default value of 0.9 during implementation as in [19].

Adam: Adaptive Moment (Adam) [19] is an optimization algorithm that merges the ideas of Momentum and RMSprop. It employs the squared gradients, like RMSprop, to adjust the learning rate and incorporates momentum by utilizing the moving average of the gradient rather than the gradient itself, similar to SGD with momentum.

$$\begin{aligned} v_{t+1} &= \beta_1 v_t - (1-\beta_1)\frac{\partial \mathcal{L}}{\partial w_t}, \\ s_{t+1} &= \beta_2 s_t - (1-\beta_2)[\frac{\partial \mathcal{L}}{\partial w_t}]^2, \\ w_{t+1} &= w_t - \alpha \frac{v_{t+1}}{\sqrt{s_{t+1}+\epsilon}} \end{aligned} \quad (11)$$

The above β_1 and β_2 hyperparameters have a default value of 0.9 in practice [19].

All the optimization algorithms discussed above are commonly applied in scene text detection and recognition [36]. However, it is worth mentioning that other optimization algorithms also are used in deep learning object detection frameworks, such as Adadelta [37], Nadam [38], AdaMax [39], and AMSGrad [40], that are not usually used for scene text detection. We examine the suitability of each optimizer or learning rate for the scene text detection framework in Section IV.

III. Methodology

In this section, we introduce the baseline and proposed module of the segmentation-based pipeline for scene text detection. Figure 2 presents an overview of the entire architecture. The network utilizes a ResNet + Feature Pyramid Network (FPN) to extract multiscale feature maps from the input images (The left part of Figure 2). These features are subsequently processed by the proposed segmentation-based attention module (SAM), which enhances the model's ability to concentrate on regions in the image that likely contain text instances during the training phase (our main contribution in this paper, which is bonded with a yellow box in the middle). The obtained features are then employed to generate the text instance masks (The right part of Figure 2).

A. Text-detection Architecture

Feature Extraction: We leverage the ResNet-based backbone network (e.g., ResNet18 or ResNet50) [20] to extract features from the input image. This paper uses two light and deep backbones for feature extraction, ResNet18 and ResNet50. The ResNet18 is mainly utilized for ablation experiments, and the ReesNet-50 is used in our final model to fairly compare it with recent methods.

FPN: We then pass the extracted features through the FPN network [41] to obtain multi-scale feature maps (1/4, 1/8, 1/16}), enabling the text detectors to be more robust in detecting text of different sizes and scales in an image. The feature maps with different scales can be written as follows:

$$F = \{f_k \epsilon \mathbb{R}^{C \times H_k \times W_k} \mid k = 0,1,2\} \quad (12)$$

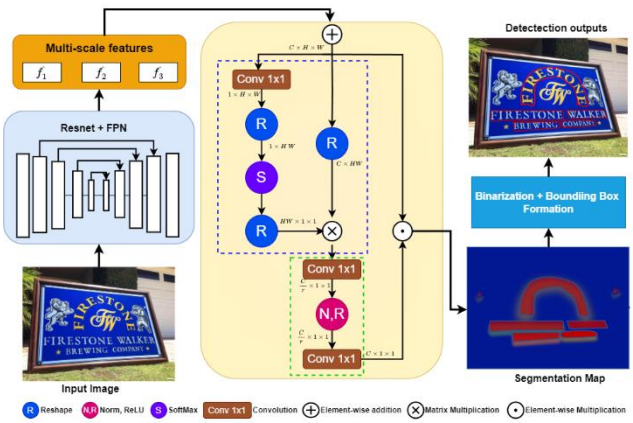


Fig. 2. The overall utilized scene text detection architecture.

The selected features are then concatenated (*concat*) and then fused by a 1x1 1D convolution layer (*conv*) as described as follows:

$$\hat{F} = Conv(Concat(f_k) \mid k = 0,1,2) \quad (13)$$

Segmentation-based Attention Module (SAM): We introduce an innovative attention module for the utilized segmentation-based architecture called SAM, as shown in Figure 2. Unlike [50, 51], SAM is designed to disseminate attention values across the spatial dimension coming from the FPN by using an element-wise multiplication strategy, which enables the model to focus more on different regions of the feature maps containing text instances.

Let $\hat{F}$ be the fused feature maps of equation (11); the output $\mathcal{S}$, which represents the global attention pooling for the SAM block (as shown with the blue dotted box in Figure 2), can be computed as follows:

$$\mathcal{S} = \hat{F}_{us} \otimes \left(\mathbb{c}_\gamma \sum_{\forall i} \frac{e^{\mathbb{c}_\alpha \hat{f}_i}}{\sum_{\forall j} e^{\mathbb{c}_\alpha \hat{f}_j}} \hat{F}_i \right) \quad (14)$$

where $\mathbb{c}_\gamma$ and $\mathbb{c}_\alpha$ are 1x1 convolutions that are learnable, $\hat{F}_{us}$ is a flattened version of vector $\hat{F}$. The j and i enumerate in the positions of all pixels. The symbol $\otimes$ represents the matrix multiplication.

Then, as depicted in the green dotted box of Figure 2, the variable $\mathbb{S}$ is transformed to capture channel-wise dependencies. The computation of this transformation can be expressed as:

$$\mathbb{S} = \sigma\left(\mathbb{c}_1\left(ReLU\left(LN(\mathbb{c}_2(\mathcal{S}))\right)\right)\right) \quad (15)$$

where σ shows the Sigmoid function. $LN(.)$ represents the Layer Normalization. $\mathbb{c}_1$ and $\mathbb{c}_2$ are 1×1 1D convolutions that are learnable.

Finally, using equations (13) and (15), the output of SAM, Γ , can be calculated as follows:

$$\Gamma = \mathbb{S} \odot \hat{F} \quad (16)$$

where $\odot$ is element-wise multiplication, the resulting outputs of equation (14) are then used to generate the segmentation map, as shown in Figure 2.

Binarization and Bounding Box Generation: After integrating the attention module into the baseline architecture, attended feature maps are obtained, as shown in Figure 2. These resulting features undergo post-processing steps to be binarized and highlight potential text regions. Using the binarized feature map, we employ differentiable binarization and bounding box label generation post-processing techniques, as described in [32, 33], to output the desired text detection. This enables the model to accurately identify and localize individual text regions within the image.

B. Loss Function

The loss function $\mathcal{L}$ is a combination of the loss for the probability map ($\mathcal{L}_s$) and the loss for the binary map ($\mathcal{L}_b$) and the loss for the threshold map $\mathcal{L}_t$ with weight factors α and β, which can be defined as follows:

$$\mathcal{L} = \mathcal{L}_{seg} + \alpha \times \mathcal{L}_b + \beta \times \mathcal{L}_t \quad (17)$$

where α and β are set to 0.9 during our experiments.

Previous scene text detection methods [21, 22, 26-28] faced challenges due to the extreme imbalance between foreground and background classes in training. Text regions occupy a small portion of the images, where the complex and straightforward examples are treated equally during training. We adopt the focal loss [35], modifying it to be more suitable for scene text detection, to address this imbalance issue. Therefore, the loss function can be defined as follows:

$$\mathcal{L}_{seg} = \mathcal{L}_b = \mathrm{FL}(\mathrm{Y}, \mathrm{Y}^*) = \frac{\sum_{i \in S_l} FL\,(p_t)}{|\mathrm{Y}^*|} = \frac{\sum_{i \in S_l} -\alpha_t (1-p_t)^{\gamma} \log p_t}{|\mathrm{Y}^*|} \quad (18)$$

In (18), Y shows the score map prediction, and Y^* illustrates the ground-truth. S_l denotes sampled set where the ratio of positives and negatives is $1:3$. The γ parameter denotes the focusing parameter, and we set it to 2 in our experiments. This parameter reduces the loss contribution from easy samples ($p_t \to 1$) and balances the straightforward and complex examples on the pixel level during optimization. We show that the leveraged loss function performs better on convergence and improves the precision performance on segmentation tasks. Furthermore, for our experiments, we use the following Binary Cross Entropy loss to compare the results of it with focal loss:

$$\mathcal{L}_{seg} = \mathcal{L}_b = \sum_{i \in S_l} Y_i \log Y_i^* + (1 - Y_i) \log (1 - Y_i^*) \quad (19)$$

C. Optimization

Which Optimizer: Section II.B addressed the availability of numerous optimization methods for training deep learning models. However, choosing an appropriate optimizer or learning rate for a specific deep-learning framework remains to be determined. Despite the widespread use of adaptive optimization techniques, especially Adam, in scene text detection, their generalization and performance beyond the training set need to be better comprehended.

Learning Rate Policy: We use different learning rate policies like step decay or constant in our experiments while training our framework. However, we show that the "poly" learning rate policy, where the learning rate is multiplied by a specific equation, is more effective than the constant learning rate or "step" learning rate (which reduces the learning rate at a fixed step size).

$$\left(1 - \frac{\text{iter}}{\text{max\_iter}}\right)^{\text{power}} \quad (20)$$

We compare the effect of this learning rate with *"step"* and *"constant"* policy in Section IV.C and IV.D.

IV. Experimental Results and Discussion

In this section, first, we provide the implementation details of the models. Next, we briefly overview the utilized benchmark datasets and evaluation metrics. Then, we conduct multiple ablation studies of the baseline and proposed schemes. Finally, the proposed technique is tested and discussed quantitatively and qualitatively on three demanding public benchmarks: MSRA-TD500 [42], Total-Text [43], and ICDAR 2015 [44]. In the end, we discuss the limitation of the proposed model and offer insights into forthcoming efforts as future work to enhance its performance.

A. Implementation Details

We employ two models for the ablation experiments and comparison with state-of-the-art scene text detection methods. The *baseline* model utilizes a ResNet18 as the feature extractor, while the final model, used for comparison, is trained with the ResNet50 backbone. The final model (the *proposed model*) is pre-trained on 200K images from the Synthetic Text dataset [45] and subsequently fine-tuned until convergence on three real-world datasets (ICDAR15, Total-Text, MSRA-TD500 - one model for each dataset).

The parameters of the utilized loss functions are optimized using the SGD+ Momentum optimizer with a poly-learning policy. Training is conducted on 2 NVIDIA Tesla A100 GPUs, with a batch size of 4, and the entire process of training and fine-tuning takes approximately 30 hours. Various augmentation techniques, such as image resizing, random rotations, and horizontal flipping, are applied during training. All models are evaluated on a single GPU of NVIDIA RTX 3080Ti with 12GB of memory.

TABLE 2.
EFFECT OF DIFFERENT OPTIMIZATION METHODS ON SCENE TEXT DETECTION MODEL.

Optimizer	Iteration	Training Loss	Val H-mean
SGD	18K	2.83	29.11
	36K	2.42	29.25
	72K	2.11	32.11
SGD+Momentum	18K	2.65	32.97
	36K	2.35	33.45
	72K	**2.09**	**35.03**
Adam	18K	4.23	27.95
	36K	3.11	28.12
	72K	2.95	28.32

B. Datasets and Evaluation Metrics

SynthText: SynthText [45] in the wild is a large synthetic dataset of 858,750 artificial scene images used to pre-train our model. These images are created by combining natural images with text generated with different fonts, sizes, orientations, and colors. The annotations include both word and character-level rotated bounding boxes and text sequences. We use this synthetic dataset to pre-train our final model to compare with other frameworks.

ICDAR2015: The ICDAR2015 [44] dataset contains 1000 images for training and 500 for testing, with annotations at the word level represented by quadrilateral boxes. This dataset presents a higher difficulty level due to its text instances with varying orientation, illumination, and complex backgrounds, and most of the images are taken in indoor environments.

Total-Text: The TotalText dataset [43], introduced in ICDAR 2017, contains 1255 training and 300 testing images with text annotations provided as polygon shapes and word-level transcriptions.

MSRA-TD500: On the other hand, the MSRA-TD500 dataset [42] consists of 300 training images and 200 test images annotated with text lines. This dataset is challenging due to multi-lingual, arbitrary-oriented, and long text lines.

Evaluation Metric: In terms of quantitative evaluation, we use the ICDAR15 Intersection over Union (IoU) Metric [44] to measure the accuracy of the detection results.

$$IOU = \frac{A(G_j \cap D_i)}{Ar(G_j \cup D_i)} \tag{21}$$

Where $A(.)$ denotes area. The IoU is then calculated by comparing the jth ground-truth bounding box to the ith detection bounding box, and detection is considered correct if the IoU value is greater than or equal to 0.5. In addition, we use the H-mean, also known as the F-score, which is a combination of precision (P) and recall (R), to measure the performance of the pre-trained model of [15, 21, 27, 34] as described in [3] and the quantitative results of recent methods in [34,52-60].

$$\text{H-mean} = 2\frac{P \times R}{P + R} \tag{22}$$

TABLE 3.
EMPLOYING DIFFERENT LEARNING HYPERPARAMETERS ON THE TOTAL-TEXT [43] VALIDATION SET.

Learning policy	Batch size	Iteration	H-mean
constant	4	50K	60.36
step	4	50K	67.48
step	4	200K	75.35
poly	4	50K	70.23
poly	4	100K	71.81
poly	6	100K	72.58
poly	**4**	**200K**	**83.26**

For evaluating the test set of Total-Text and MSRA-TD500 datasets, we utilize similar evaluation metrics introduced in [43] and [42], respectively.

C. Effect Of Different Optimizers

To show the effect of the different optimizers, we first trained the baseline framework with ResNet18 backbone and without the SAM block on the 70% of SynthText dataset [47] for 80k iterations with a fixed learning rate of 1×10^{-4} in three SGD, SGD+Momentum (0.9), and Adam optimization algorithms. Adam works well in practice and outperforms other adaptive techniques, and it is the default optimizer in many deep learning-based detectors. Then, we evaluated our pre-trained models with three optimizers on the remaining 30% of the SynthText dataset as validation, as shown in Table 2. Although more epochs of training, we observed more decline in loss of the SGD optimizer compared to Adam in higher iterations.

Generalizability is a crucial characteristic in many deep learning models, which shows how a trained model on one dataset is capable of detecting challenging text in the validation set or other datasets; As can be seen from Table 2., SGD and SGD+Momentum provide better validation H-mean performance compared to Adam as the number of iteration increased. For instance, in the 72K iteration, SGD+Momentum achieved about 35%, while Adam remained at about 28%. Therefore, SGD has better generalization capability than Adam for our proposed method, and for the rest of this paper, we used SGD+Momentum to train our models.

D. Effect of Poly Learning Rate

Choosing a proper learning rate is also essential during the training of deep learning networks. For this purpose, we experimented with comparing the effect of different learning rates on our scene text detection problem. Table 3 illustrates the H-mean performance of employing "constant," "step," and "poly" learning rates. We used a fixed learning rate of 1×10^{-4} during training for constant policy. As shown in

Table 3, employing "a poly" learning rate of (20) (with $power = 0.9$) and using the same batch size and same training

TABLE 4.
ABLATION STUDY TO EVALUATE THE EFFECT OF THE DESINED SAM BLOCK BY USING DIFFERENT BACKBONE AND LOSS FUNCTIONS ON THE TOTAL-TEXT [43] VALIDATION SET.

Model	Backbone	Attention module	Loss function	Iteration	H
Baseline	ResNet18	--	BCE	200K	79.3
Baseline	ResNet18	--	FL	200K	83.4
Proposed	ResNet18	SAM	FL	200K	84.9
Proposed	ResNet50	SAM	FL	200K	**86.5**

iterations yields 2.75% and 9.87% better performance than employing "step" and "constant" learning rate policy, respectively. The "constant" learning rate does not improve the H-mean performance of the model, and significantly has la ower performance in terms of H-mean from other learning rate policies.Therefore, we ignored it in the rest of our experiments. Fixing the batch size and increasing the training iteration to 100K improves the performance to 71.81% (1.58↑);

By increasing the batch size and the same iteration of 50K, we still experienced increased H-mean performance. We then reduced the batch size to 4 and found that comparable performance remains (71.81% versus 72.58%). Ultimately, we applied a batch size of 4 and 200K iterations to maintain similar training iterations as a "step" policy. We observed a significant H-mean performance of 83.26 (7.91% improvement over the "step" policy).

E. Effect of the Designed SAM, Feature Extraction Backbone, and Loss Functions

We conducted the following experiments on the benchmark Total-Text dataset to see the effect of the designed SAM block, different feature extraction backbones, and loss functions in the baseline and proposed model, shown in Table 4. We employed the "poly" learning policy and SGD+Momentom optimizer in all these experiments and set the batch size to 4.

We first eliminated the SAM and used the ResNet18 backbone to compare the effect of FL in (18) and BCE loss in (19). It can be seen from Table 4 that the model with focal loss outperformed the model with BCE loss using a similar setup of the model because this loss is more robust in challenging cases like the vertical or different scale of the text and designed to strike the imbalance issue of positive or negative text existing in the scene image. Then by applying the SAM to the architecture, as seen from Table 4, the H-mean performance boosted on a similar setup by ~1.5%, which confirms its effectiveness in the utilized architecture.

Recent methods in scene text detection [53-60] use ResNet50 instead of ResNet18 due to its deeper architecture and increased model capacity, enabling better feature representation and performance for handling complex and

TABLE 5.
QUANTITATIVE COMPARISON OF THE PROPOSED MODEL AMONG SOME OF THE RECENT METHODS ON TOTAL-TEXT DATASET.

Method	Total-Text			FPS
	Precision	Recall	H-mean	
CRAFT [15]	87.6	79.9	83.6	6.00
PSENet [34]	84.0	78.0	80.9	2.70
TextSnake [28]	82.7	74.5	78.4	1.10
DBNet [33]	87.1	82.5	84.7	19.0
DRRG [54]	86.5	84.9	85.7	3.50
FCENet [55]	89.3	82.5	85.8	–
PCR [56]	88.5	82.0	85.2	10.7
TextBPN [57]	90.3	84.7	87.4	10.6
TextDCT[53]	87.2	82.7	84.9	–
ABCNetV2[58]	89.2	84.1	87.0	10.0
FSG [52]	90.7	**85.7**	**88.1**	13.1
TPSNet [59]	89.2	85.0	86.6	11.6
LeafText [60]	**90.8**	84.0	87.3	–
Baseline	88.8	78.7	83.4	**31.0**
Proposed Model	90.5	82.9	86.5	**22.0**

diverse visual patterns in text regions. We use the ResNet50 as our final feature extraction backbone to compare the recent models fairly. As shown, ResNet50 improved the detection performance of the model in terms of H-mean by ~1.5% compared to using ResNet18 because it allows the model to capture more complex and high-level features from the input images. The deeper architecture of ResNet50 provides a larger receptive field, which means that the network can consider a broader context of the input image.

F. Comparison with State-of-the-Art Text Detection Methods

We also compare our proposed model with several state-of-the-art scene text detectors [15, 28, 33, 34, 52-60] in Table 5 and Table 6 across three benchmark datasets: Total-Text, ICDAR15 [44], and MSRA-TD500 [42]. The proposed model utilizes a ResNet50 and the designed SAM block. The baseline model leverages a ResNet18 in its backbone as a feature extractor. Despite our final proposed model being pre-trained on a subset of synthetic dataset images and featuring a straightforward design, it achieved competitive performance in accuracy and efficiency, outperforming many recent state-of-the-art methods that often undergo pretraining on multiple datasets and involve complex design and post-processing.

The results in Table 5 demonstrate that the proposed model significantly outperforms the baseline model on the Total-Text dataset, which consists of many challenging text images, including curved and irregular text instances. The effectiveness of the SAM module is particularly evident in its

recall performance, improving by approximately 4% compared to the baseline.

TABLE 6.
QUANTITATIVE COMPARISON OF THE PROPOSED MODEL AMONG SOME OF THE RECENT TEXT DETECTION METHODS ON ICDAR15 AND MSRA-TD500 DATASETS.

Method	ICDAR15			MSRA-TD500		
	Precision	Recall	H-mean	Precision	Recall	H-mean
EAST [21]	83.6	73.5	78.2	–	–	–
Pixellink [27]	82.8	81.6	82.2	83.0	73.2	77.8
CRAFT [15]	82.2	77.8	82.2	88.2	78.2	82.9
PSENet [34]	86.9	84.5	85.7	–	–	–
TextSnake [28]	84.9	80.4	82.6	83.2	73.9	78.3
DBNet [33]	86.8	78.4	82.3	91.5	79.2	84.9
DRRG [54]	88.5	84.7	86.6	88.0	82.3	85.0
FCENet [55]	90.1	82.6	86.2	–	–	–
PCR [56]	–	–	–	87.6	77.8	82.4
TextBPN [57]	–	–	–	86.6	84.5	85.6
TextDCT[53]	88.9	84.8	86.8	–	–	–
ABCNetV2[58]	90.4	86.0	88.1	89.4	81.3	85.2
FSG [52]	**90.9**	**87.3**	**89.1**	91.6	**84.8**	**88.1**
TPSNet [59]	90.5	85.1	87.7	–	–	–
LeafText [60]	88.9	82.3	86.1	**92.1**	83.8	87.8
Baseline	87.8	77.6	82.4	87.6	79.4	83.3
Proposed Model	90.3	81.3	85.6	89.8	82.4	85.9

As indicated in Table 5, the final model also demonstrated excellent performance in terms of Frames per second (FPS), achieving a rate of 22 FPS during inference, surpassing several other detectors. This high FPS capability highlights the applicability of our model in real-time detection applications. It is also worth mentioning that the designed SAM has minimal impact on the model's inference time.

As shown in Table 6, the designed SAM block and ResNet50 feature extractor boosted the performance of the proposed model by ~3% and ~2.5% in terms of H-mean performance on the ICDAR15 and MSRA-TD500 datasets, respectively, which are explicitly designed for multi-oriented scene text detection. The well performances on these datasets, confirm the proposed model's capability in detecting challenging text instances that are abundant in ICDAR15 and MSRA-TD500 datasets.

Finally, we present qualitative results in Figure 3, showcasing the detection performance of our proposed model on various samples from the Total-Text, ICDAR15, and MSRA-TD500 datasets. The results demonstrate the model's effectiveness in handling challenging cases, including multi-oriented, vertical, and curved text, small and complex fonts, and different languages, indicating its robustness and versatility across diverse scene text scenarios.

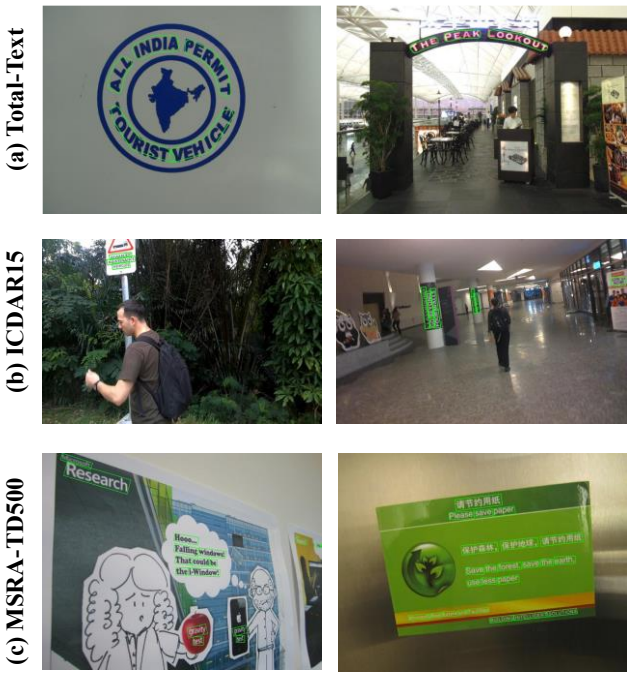


Fig. 3. Qualitative results of our final proposed method on benchmark datasets: (a) Total-Text [43], (b) ICDAR15 [44], and (c) MSRA-TD500 [42]. Detected text instances are shown with green color.

G. Limitation and Future Work for improvement

The presented model exhibits a significant limitation in cases where two-word instances in the image are closely positioned (i.e. when the first word's last character is close to the second word's first character). This scenario and the detection of crossed-word instances pose challenges and remain an unsolved problem in segmentation-based text detection architectures. Additionally, the proposed model encounters difficulty accurately detecting text instances under severe low resolution, low contrast, and occlusion conditions.

To address these limitations, the following suggested future work may enhance the model's ability to detect cha text instances. (1) One promising avenue for future work is to leverage recent advancements in natural language processing algorithms. Specifically, integrating pre-trained language model modules such as Generative Pre-training Transformer (GPT) [62] into the text detection and recognition framework could prove beneficial. By employing compositionality techniques [63, 64] alongside the language model, the model may be better equipped to infer and reconstruct the uncaptured characters in the detected text instances. (2) Another potential area for future research is exploring context-aware object

interaction.

By incorporating contextual information from the surrounding objects and the overall scene, the model could better understand the text instances' spatial relationships and textual coherence. This context-aware approach might enhance the model's ability to detect accurately and segment text instances closely intertwined or obscured by visual obstacles. (3) Given the challenges posed by low resolution, low contrast, and occlusion, another future direction is to explore adversarial training techniques. It could become more robust against various types of image degradation and perturbations by incorporating adversarial examples during the model's training process. This enhanced robustness might improve the model's performance in adverse real-world conditions, leading to more reliable text detection results.

V. Conclusions

This paper introduced a simple yet effective segmentation-based architecture for detecting text of arbitrary shapes in wild images. We thoroughly investigated deep-learning optimizers and loss functions applicable to our baseline pipeline. Additionally, we designed a specific attention block, the segmentation-based attention module (SAM), to focus on regions more likely to contain text instances. Through extensive ablation studies, we assessed the impact of the proposed module and analyzed different optimization algorithms and loss functions. Our experiments demonstrated that the SGD + Momentum optimizer with a "poly" learning rate policy offers superior detection performance, and the focal loss-based loss function proves more robust than binary cross-entropy loss. Comparing our model to recent state-of-the-art scene text detection methods on three benchmark datasets, our final model with the designed SAM significantly outperforms the baseline model and achieves competitive detection performance and efficiency. We also discussed the limitations of our architecture and outlined future research directions to further enhance text detection capabilities.

REFERENCES

[1] H. Lin, P. Yang, and F. Zhang, "Review of scene text detection and recognition, " *Archives of Computational Methods in Eng*., pp. 1–22, 2019.

[2] X. Liu, G. Meng, and C. Pan, "Scene text detection and recognition with advances in deep learning: A survey," *Int. J. on Document Anal. and Recognition (IJDAR)*, pp. 1–20, 2019.

[3] Z. Raisi, M. A. Naiel, G. Younes, P. Fieguth, and J. Zelek, "Smart text reader system for people who are blind using machine and deep learning, " *Machine Learning Algorithms for Signal and Image Processing*, pp. 161–200, 2022.

[4] Z. Raisi and J. Zelek, "Text detection & recognition in the wild for robot localization," *arXiv preprint arXiv:2205.08565*, 2022.

[5] Z. Raisi, M. A. Naiel, P. Fieguth, S. Wardell, and J. Zelek, "Text detection and recognition in the wild: A review, " arXiv preprint arXiv:2006.04305, 2020.

[6] Z. Raisi, "Text detection and recognition in the wild, " 2022. [7] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," Image and vision computing, vol. 22, no. 10, pp. 761–767, 2004.

[8] L. Neumann and J. Matas, "A method for text localization and recognition in real-world images, " in Proceedings Asian Conference on Computer Vision. Springer, 2010, pp. 770–783.

[9] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 2963–2970.

[10] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector, " in Eur. Conference on Computer Vision. Springer, 2016, pp. 21–37.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779– 788.

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Proceedings Adv. in Neural Info. Process. Sys., 2015, pp. 91–99.

[13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[14] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in Proceedings IEEE Int. Conference on Computer Vision, 2017, pp. 2961–2969.

[15] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, "Character region awareness for text detection," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2019.

[16] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning, " in International conference on machine learning, 2013, pp. 1139– 1147.

[17] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," Journal of machine learning research, vol. 12, no. Jul, pp. 2121–2159, 2011.

[18] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude," COURSERA: Neural networks for machine learning, vol. 4, no. 2, pp. 26–31, 2012.

[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition, " Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2015.

[21] X. Zhou, C. Yao, H. Wen, Y. Wang, S. Zhou, W. He, and J. Liang, "East: an efficient and accurate scene text detector, " in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5551–5560.

[22] M. Liao, B. Shi, X. Bai, X. Wang, and W. Liu, "Textboxes: A fast text detector with a single deep neural network, " in Proceedings AAAI Conference on Artificial Intelligence, 2017.

[23] Y. Liu and L. Jin, "Deep matching prior network: Toward tighter multi-oriented text detection," in Proceedings IEEE

Conference on Computer Vision and Pattern Recognition, 2017, pp. 1962–1969.
[24] M. Liao, B. Shi, and X. Bai, "Textboxes++: A single-shot oriented scene text detector, " IEEE Trans. on Image Process. , vol. 27, no. 8, pp. 3676–3690, 2018.
[25] J. Ma, W. Shao, H. Ye, L. Wang, H. Wang, Y. Zheng, and X. Xue, "Arbitrary-oriented scene text detection via rotation proposals, "IEEE Trans. on Multimedia", vol. 20, no. 11, pp. 3111-3122, 2018.
[26] Z. Zhang, C. Zhang, W. Shen, C. Yao, W. Liu, and X. Bai, "Multi-oriented text detection with fully convolutional networks," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4159–4167.
[27] D. Deng, H. Liu, X. Li, and D. Cai, "Pixellink: Detecting scene text via instance segmentation," in Proceedings AAAI Conference on Artificial Intelligence, 2018.
[28] S. Long, J. Ruan, W. Zhang, X. He, W. Wu, and C. Yao, "Textsnake: A flexible representation for detecting text of ar-arbitrary shapes," in Proceedings Eur. Conference on Computer Vision (ECCV), 2018, pp. 20–36.
[29] C. Yao, X. Bai, N. Sang, X. Zhou, S. Zhou, and Z. Cao, "Scene text detection via holistic, multi-channel prediction," arXiv preprint arXiv:1606.09002, 2016.
[30] B. Shi, X. Bai, and S. Belongie, "Detecting oriented text in natural images by linking segments, " in Proceedings *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2550–2558.
[31] W. Wang, E. Xie, X. Song, Y. Zang, W. Wang, T. Lu, G. Yu, and C. Shen, "Efficient and accurate arbitrary-shaped text detection with pixel aggregation network, " in Proceedings of the IEEE Int. Conference on Computer Vision, 2019, pp. 8440–8449.
[32] P. Yang, G. Yang, X. Gong, P. Wu, X. Han, J. Wu, and C. Chen, "Instance segmentation network with self-distillation for scene text detection," IEEE Access, vol. 8, pp. 45 825–45 836, 2020.
[33] M. Liao, Z. Wan, C. Yao, K. Chen, and X. Bai. Real-time scene text detection with differentiable binarization. In AAAI Conference on Artificial Intelligence, pages 11474–11481, 2020.
[34] W. Wang, E. Xie, X. Li, W. Hou, T. Lu, G. Yu, and S. Shao, "Shape robust text detection with progressive scale expansion network, " arXiv preprint arXiv:1903.12473, 2019.
[35] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in Proceedings IEEE Int. Conference on Computer Vision, 2017, pp. 2980–2988.
[36] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht, "The marginal value of adaptive gradient methods in machine learning, " in Advances in Neural Information Processing Systems, 2017, pp. 4148–4158.
[37] M. D. Zeiler, "Adadelta: an adaptive learning rate method," arXiv preprint arXiv:1212.5701, 2012.
[38] T. Dozat, "Incorporating nesterov momentum into adam," 2016
[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014
[40] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of Adam and beyond, " arXiv preprint arXiv:1904.09237, 2019.
[41] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), July 2017.
[42] C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, "Detecting texts of arbitrary orientations in natural images, " in IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012, pp. 1083–1090.
[43] C. K. Ch'ng and C. S. Chan, "Total-text: A comprehensive dataset for scene text detection and recognition, " in Proceedings IAPR Int. Conf. on Document Anal. and Recognition (ICDAR), vol. 1, 2017, pp. 935–942.
[44] D. Karatzas, L. Gomez-Bigorda, A. Nicolaou, S. Ghosh, A. Bagdanov, M. Iwamura, J. Matas, L. Neumann, V. R. Chandrasekhar, S. Lu et al., "Icdar 2015 competition on robust reading," in Proceedings Int. Conf. on Document Anal. and Recognition (ICDAR), 2015, pp. 1156–1160.
[45] A. Gupta, A. Vedaldi, and A. Zisserman, "Synthetic data for text localization in natural images, " in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2016, last retrieved March 11, 2020.
[46] J. Liu, X. Liu, J. Sheng, D. Liang, X. Li, and Q. Liu, "Pyramid mask text detector, " CoRR, vol. abs/1903.11800, 2019.
[47] A. Gupta, A. Vedaldi, and A. Zisserman, "Synthetic data for text localization in natural images, " in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2315–2324.
[48] N. Darjani and H. Omranpour, "Comprehensive learning polynomial auto-regressive model based on optimization with application of time series forecasting," International Journal of Industrial Electronics Control and Optimization, vol. 5,no. 1, pp. 43–50, 2022.
[49] S. Kalantari, M. Ramezani, and A. Madadi, "Introducing a new hybrid adaptive local optimal low-rank approximation method for denoising images," International Journal of Industrial Electronics Control and Optimization, vol. 3, no. 2, pp. 173–185, 2020.
[50] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 13 713–13 722.
[51] Y. Cao, J. Xu, S. Lin, F. Wei, and H. Hu, "Gcnet: Non-local networks meet squeeze-excitation networks and beyond," in Proceedings of the IEEE/CVF international conference on Computer vision Workshops, 2019.
[52] J. Tang, W. Zhang, H. Liu, M. Yang, B. Jiang, G. Hu, and X. Bai, "Few could be better than all: Feature sampling and grouping for scene text detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 4563–4572.
[53] Y. Su, Z. Shao, Y. Zhou, F. Meng, H. Zhu, B. Liu, and R. Yao, "Textdct: Arbitrary-shaped text detection via discrete cosine transform mask," IEEE Transactions on Multimedia, 2022.
[54] S.-X. Zhang, X. Zhu, J.-B. Hou, C. Liu, C. Yang, H. Wang, and X.-C. Yin, "Deep relational reasoning graph network for arbitrary shape text detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 9699–9708.
[55] Y. Zhu, J. Chen, L. Liang, Z. Kuang, L. Jin, and W. Zhang, "Fourier contour embedding for arbitrary-shaped text detection," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 3123–3131.
[56] P. Dai, S. Zhang, H. Zhang, and X. Cao, "Progressive contour regression for arbitrary-shape scene text detection," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 7393–7402.
[57] S.-X. Zhang, X. Zhu, C. Yang, H. Wang, and X.-C. Yin, "Adaptive boundary proposal network for arbitrary shape text detection," in Proceedings of the IEEE/CVF International Conference on computer vision, 2021, pp. 1305–1314.

[58] Y. Liu, C. Shen, L. Jin, T. He, P. Chen, C. Liu, and H. Chen, "Abcnet v2: Adaptive bezier-curve network for real-time end-to-end text spotting," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 11, pp. 8048–8064, 2021.
[59] W. Wang, Y. Zhou, J. Lv, D. Wu, G. Zhao, N. Jiang, and W. Wang, "Tpsnet: Reverse thinking of thin plate splines for arbitrary shape scene text representation," in Proceedings of the 30th ACM International Conference on Multimedia, 2022, pp. 5014–5025.
[60] C. Yang, M. Chen, Y. Yuan, and Q. Wang, "Text growing on leaf," IEEE Transactions on Multimedia, 2023.
[61] V. Nazarzehi and R. Damani, "Decentralised optimal deployment of mobile underwater sensors for covering layers of the ocean," Indonesian Journal of Electrical Engineering and Computer Science, vol. 25, no. 2, pp. 840–846, 2022.
[62] D. M. Katz, M. J. Bommarito, S. Gao, and P. Arredondo, "Gpt-4 passes the bar exam," Available at SSRN 4389233, 2023.
[63] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," Behavioral and brain sciences, vol. 40, p. e253, 2017.
[64] Z. Raisi and J. Zelek, "Occluded text detection and recognition in the wild," in 2022 19th Conference on Robots and Vision (CRV). IEEE, 2022, pp. 140–150.

Zobeir Raisi was born in Chabahar, Iran in 1987. He received his Ph.D. degree in 2022 from the Vision Image Processing Lab (VIPLab) at Systems Design Engineering Departments, University of Waterloo, Waterloo, Ontario, Canada. Currently, he is an assistant professor in the Department of Electrical Engineering, Chabahar Maritime University, Iran. His research interests include computer vision, artificial intelligence, robotics, and image processing.

John Zelek received his Ph.D. degree in philosophy of electrical engineering from the Centre for Intelligent Machines (CIM), McGill University, Montreal, QC, Canada, in 1996. He is currently a Professor of the Systems Design Engineering Department, University of Waterloo, Waterloo, ON, Canada and the Co-Director of the Vision Image Processing (VIP) Laboratory, University of Waterloo, ON, Canada. He has published over 300 refereed articles and has been a co-founder of five different startup companies from the University of Waterloo and has been an advisor for various other companies. His research interests include computer vision, AI, robotics, infrastructure monitoring, autonomous vehicles, image processing, augmented reality, assistive technology, to name a few.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Adaptive Input-Output Feedback Linearization Control for Islanded Inverter-Based Microgrids

Navid Reza Abjadi[1]

Faculty of Engineering, Shahrekord University. Shahrekord, Iran.[1]
Corresponding author's email: abjadi.navidreza@sku.ac.ir, navidabjadi@yahoo.com

Article Info

Article type:
Research Article

Article history:
Received: 16-February-2023
Received in revised form: 21-July-2023
Accepted: 16-Aug-2023
Published online: 25-Aug-2023

Keywords:
Adaptive control,
Distributed generation,
Feedback linearization,
Master-slave strategy,
Microgrid.

ABSTRACT

Due to the growth of renewable energies and the need for sustainable electrical energy, AC microgrids (MGs) have been the subject of intense research. Medium voltage MGs will soon have a special place in the power industry. This paper uses a new and effective control scheme for islanded inverter-based medium voltage MGs using the master-slave (MS) technique. The controllers only need local measurements. The designed controls are based on adaptive input-output feedback linearization control (AIOFLC). These controls have a high-performance response, and are robust against some uncertainties and disturbances. The use of the designed control scheme makes the output voltage of distributed generation (DG) sources have negligible harmonics. Besides, the generated voltage and active/reactive powers track their references effectively. The model of the inverter-based DGs is considered in a stationary reference frame, and there is no need for any coordinate frame transformation. The control method presented in this paper can be used for MGs with any number of inverter-based DGs and parallel inverters. The effectiveness of the proposed control scheme is evaluated by simulation in SIMULINK/MATLAB environment and compared with that of feedback linearization control (FLC) and conventional sliding mode control (CSMC).

I. Introduction

Today, the severe warming of the earth due to greenhouse gases and environmental pollution endangered human lives in some areas, leading to the desire of many countries to be independent of fossil fuels. Accordingly, many efforts have been made to produce electricity from renewable energies such as wind and sun. One of the best ways to exploit these energy sources is to use them in the MG. The MG may be connected to the grid or disconnected from the grid. When MG is disconnected from the grid, it is called an islanded MG, wherein the voltage and frequency must be controlled and the required power must be divided between the DG sources. An islanded MG is a complicated multivariable system with difficult high performance control of voltage amplitude, frequency and active/reactive powers.

In an islanded MG, two types of strategies to share active/reactive powers between DGs are communication-free and communication-based strategies [1]. The droop control strategy is known as a communication-free strategy because it

requires local measurements [2]; however, this strategy is based on small signal models, which may cause instability in large signal changes. In addition, it causes DC bus voltage deviation [3]. In [4], proportional resonance (PR) control with a droop control strategy is used for an islanded MG. PR control is not robust against uncertainties, disturbances, and operating point changes. The well-known communication-based strategies are the centralized control strategy, distributed control strategy, and MS strategy [5]. Centralized control has a very good dynamic response but needs all of the information and measurements of the DGs, which may be far from each other such that the whole system may fail with a small defect [6]. In addition, it requires a high-speed expensive telecommunication. A low bandwidth link is needed for distributed control and MS strategy; since they need less MG data. MS strategy can result in excellent power sharing and is easily implemented [1].

In the MS strategy, the DG with the largest power capacity is called the master unit (MU), and the other DGs are called the slave units (SU). The MU controls the voltage and frequency of the MG, and the SUs control their active and reactive power output [1]. In [7], power sharing in MS strategy is investigated. In [8], a two-level control for MS strategy is designed which does not have a good dynamic response and does not work well in the face of disturbances and uncertainties. In [1], by adopting the MS strategy and proportional-integral (PI) control, the result of communication delay is investigated for two parallel inverters. Tuning the PI controllers depends on the operating points, and the method is not robust. In [9], a communication-free MS strategy is proposed for an MG. In this strategy, the MU is a synchronous generator and the SUs are current source inverters (CSIs).

In [10], the effect of coefficients of conventional decoupling PI in MS strategy is studied. In [11], a cascaded PI-based control is analyzed using the MS strategy in islanded MG. In [12], μ synthesis is used for MG with MS strategy. The DGs in [12] use a three-stage solid-state transformer with a complicated structure. In [13], adopting MS strategy, an H_∞ controller is proposed for the MU. However, for the SUs, PI controllers are used in the dq reference frame. Therefore, the overall closed-loop system is not robust. In addition, the design procedure of the H_∞ controller has several steps, and matrix manipulations are needed.

In [14], a nonlinear backstepping control is designed for an MG with MS strategy, which needs loads measurements. Backstepping control in [14] is a recursive method with many steps and is not robust.

CSMC is a method that has attracted such attention regarding its simple implementation and robustness. The main problem of CSMC is the chattering phenomenon. In [15]. adaptive SMC and adaptive feedback linearization methods are used for an MG with MS strategy. Given that there is no guarantee for the convergence of estimates to real values, the amount of control effort can be high, and instability may occur. In addition, the lumped uncertainties considered in the proposed control design in [16] are considered constants, while they are not constant. In fact, these lumped uncertainties depend on voltages, currents, and parameters uncertainties. In [16], the voltages and currents of an MG are controlled using high-order SMC methods by accepting centralized control strategy. The control scheme proposed in [16] is rather complex and it must be redesigned with the change of MG completely.

In [17], a secondary control method is used for a hybrid droop and MS structure. The control scheme of [17] requires internal current control loops, several abc/dq conversions, and too many parameters. In addition, it is not robust.

Model predictive control design for MS inverters is presented in [18]. In this paper, it is claimed that the controllers are robust with respect to uncertainties, while the obtained models are dependent on system parameters to predict voltages and currents.

A new enhanced droop control for MGs based on small signal models is proposed in [19]. Large disturbances and variations may cause instability in the overall proposed system.

In [20], an enhanced state feedback for secondary control in an islanded MG is designed. The control method of [20] is not robust with respect to MG parameters uncertainties.

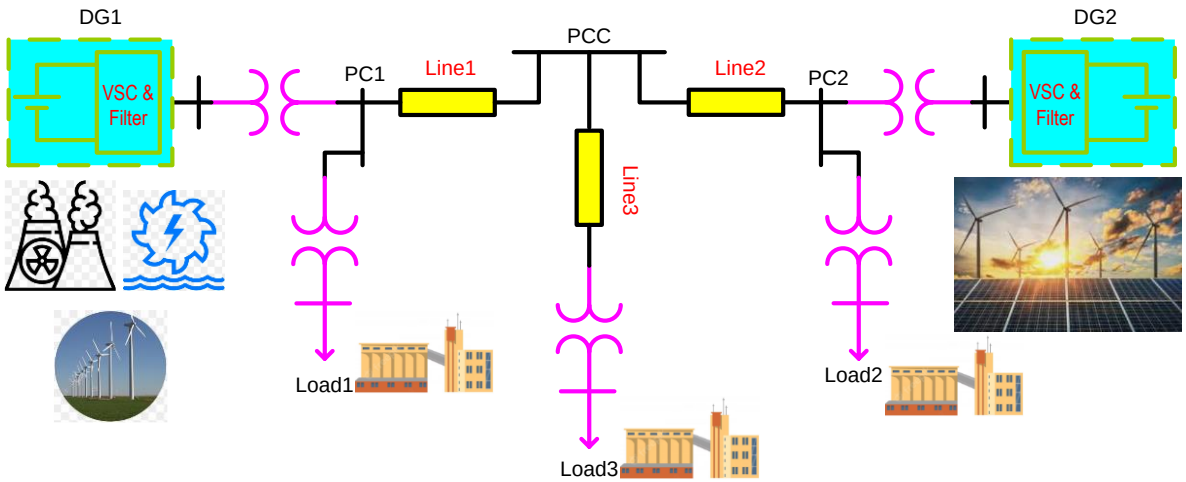
Nonsingular terminal Sliding mode control (TSMC) is used for Islanded inverter-based MGs in [21]. A high-performance control scheme is achieved in [21]; however, fractional order calculation in TSMC causes difficulties in real time implementation especially when low-cost microcontrollers are used.

Adaptive feedback linearization control is a robust control. In [22], an AIOFLC is designed for a grid-connected inverter with an L filter to control the active and reactive powers injected into the grid.

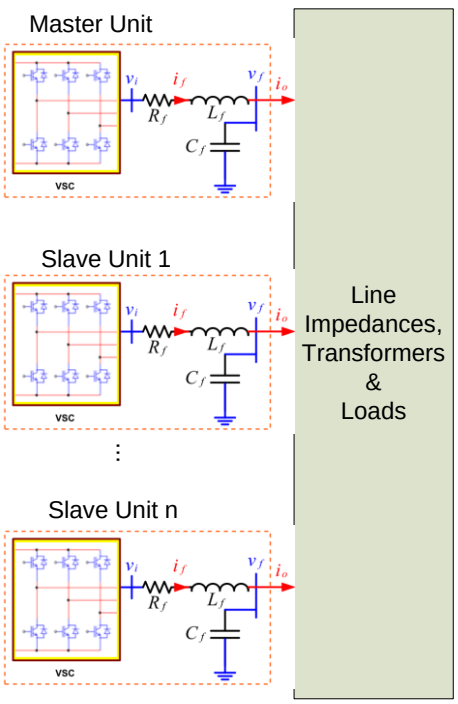
This paper designs novel controls for inverter-based DGs of islanded MGs considering MS strategy. The inverters filters are LC filters and the applied control technique is AIOFLC. The modeling equations of MU and SUs are rewritten to suitable forms for AIOFLC. An AIOFLC is designed for MU to control the voltage of the MU bus. Another AIOFLC is designed for SUs to control the active and reactive powers delivered by SUs. There is symmetry in the control design because the AIOFLCs are designed for both master and slave units. Using the proposed control method, a high-performance closed–loop control system is achieved. The stability is guaranteed. The microgrid handles the load switching, uncertainties and disturbances properly. Several case studies are investigated and the simulation results are represented.

II. MG configuration and modeling

The islanded MG investigated in this paper is given in Fig 1a.



(a)



(b)

Fig. 1. Islanded MG and MS strategy, (a) A MG with two inverter-based DGs, (b) MS strategy block diagram

This MG has two inverter-based DGs and three loads; however, the models, power sharing strategy, and the designed controls can be used for any islanded MG with any number of inverter-based DGs. The DG with the larger power capacity is selected as the MU and the other as the SU. MS block diagram is illustrated in Fig. 1b. Using MS strategy, the voltage (amplitude, and frequency) is regulated by MU. In fact, MU is a grid forming inverter and it should be the largest DG. Therefore, it can obtain its energy from any source including, photovoltaic. However, to achieve maximum power point tracking (MPPT), it's better to use photovoltaic power sources as SUs (grid following units), because the active/reactive powers of SUs are controlled and MPPT can be implemented easily.

Fig. 2 illustrates an inverter-based DG unit connected to an LC filter. The state space equations in stationary reference frame for the system shown in Fig. 2 can be written as follows [15]

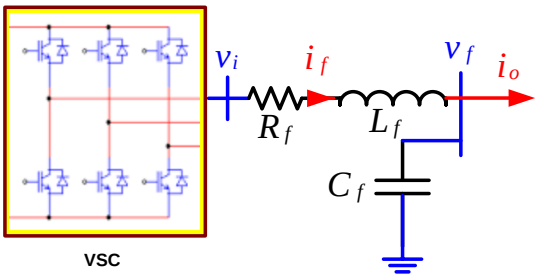


Fig. 2. An inverter-based DG unit

$$di_f/dt = a(v_i - bi_f - v_f) + \xi_i \quad (1)$$

$$dv_f/dt = c(i_f - i_o) + \xi_v \quad (2)$$

where $i_f = [i_{f\alpha}\ i_{f\beta}]^T$, $i_o = [i_{o\alpha}\ i_{o\beta}]^T$, $v_i = [v_{i\alpha}\ v_{i\beta}]^T$, $v_f = [v_{f\alpha}\ v_{f\beta}]^T$, $a = 1/L_f$, $b = R_f$ and $c = 1/C_f$. $\xi_i = [\xi_{i\alpha}\ \xi_{i\beta}]^T$, and $\xi_v = [\xi_{v\alpha}\ \xi_{v\beta}]^T$ are lumped uncertainties:

$$\xi_i = \Delta a(v_i - bi_f - v_f) - (a + \Delta a)\Delta b i_f \quad (3)$$

$$\xi_v = \Delta c(i_f - i_o) \quad (4)$$

where Δa, Δb and Δc are parameters uncertainties.

The powers injected into the filter bus are obtained as [15]

$$P = \frac{3}{2}(v_{f\alpha}i_{f\alpha} + v_{f\beta}i_{f\beta}) \quad (5)$$

$$Q = \frac{3}{2}(v_{f\beta}i_{f\alpha} - v_{f\alpha}i_{f\beta}) \quad (6)$$

The active and reactive powers are nonlinear functions of the state variables.

III. Controllers design

In this section, first, the design of an AIOFLC for a class of nonlinear systems based on the Lyapunov function is discussed, followed by employing this control design for MU and SUs.

A. AIOFLC design

Consider the following nonlinear system [22],

$$\dot{y} = f(x) + q\theta + g(x)u \quad (7)$$

where y is the output; x is states vector; $f(x)$, $g(x)$ and q are known vectors; u is control input; θ is the difference between actual and nominal values of parameters.

The following Lyapunov function is nominated [22],

$$V = \frac{1}{2}y^2 + \frac{1}{2}\tilde{\theta}^{\mathrm{T}}\Gamma\tilde{\theta} \quad (8)$$

where $\tilde{\theta} = \theta - \hat{\theta}$, and $\hat{\theta}$ is the estimated vector of θ; Γ is a strictly positive definite constant matrix.

Taking the time derivative of (6) and using (5), gives

$$\dot{V} = y\dot{y} + \tilde{\theta}^{\mathrm{T}}\Gamma\left(-\dot{\hat{\theta}}\right) = y(f(x) + q\theta + g(x)u) + \tilde{\theta}^{\mathrm{T}}\Gamma(-\dot{\hat{\theta}}) = y(f(x) + q\hat{\theta} + g(x)u) + \tilde{\theta}^{\mathrm{T}}(yq^{\mathrm{T}} - \Gamma\dot{\hat{\theta}}) \quad (9)$$

Considering the following control law and adaptation law, $\dot{V}$ in (7) becomes negative semi-definite [22].

$$u = \frac{1}{g(x)}(-f(x) - q\hat{\theta} - ky) \quad (10)$$

$$\dot{\hat{\theta}} = \Gamma^{-1}yq^{\mathrm{T}} \quad (11)$$

The positive definiteness of V and negative semi definiteness of $\dot{V}$ yield that y and $\tilde{\theta}$ are bounded. A sufficient condition for convergence of y to 0 is the uniformly continuousness of $\dot{V}$, which is satisfied if $\ddot{V}$ is bounded [23].

Substituting (10) and (11) in (9), yields

$$\dot{V} = -ky^2 \quad (12)$$

Taking the time derivative of (12) and using (7), gives

$$\ddot{V} = -2ky\dot{y} = -2ky(f(x) + q\theta + g(x)u) \quad (13)$$

Since y and $\tilde{\theta}$ are bounded, all of the variables and functions in right hand side of (13) are bounded. Therefore, $\ddot{V} < \infty$, which yields the convergence of y to 0.

B. AIOFLC for MU

To achieve a robust high-performance control of MU bus voltage, the AIOFLC is designed for MU. For this purpose, the dynamical equation of MU should be rewritten in the canonical form given in (7). Consider the following output

$$y = (v_f - v_f^*) + k_1(\dot{v}_f - \dot{v}_f^*) \quad (14)$$

where v_f^* is the MU bus voltage reference.

Taking the time derivative of (14) and using (1)-(4), gives

$$\dot{y} = (\dot{v}_f - \dot{v}_f^*) + k_1(\ddot{v}_f - \ddot{v}_f^*) = c(i_f - i_o) + \xi_v + k_1 c[a(v_i - bi_f - v_f) + \xi_i] - k_1 c\frac{di_o}{dt} - k_1\dot{\xi}_v - \dot{v}_f^* - k_1\ddot{v}_f^* \quad (15)$$

where the time derivative of ξ_v is given by

$$\dot{\xi}_v = \Delta c\left(\frac{di_f}{dt} - \frac{di_o}{dt}\right) \quad (16)$$

Substituting for $\dot{\xi}_v$ in (15), (15) can be considered in the canonical form (7) with

$$f(x) = c(i_f - i_o) + k_1 ca(bi_f + v_f) - k_1 c\frac{di_o}{dt} - \dot{v}_f^* - k_1\ddot{v}_f^* \quad (17)$$

$$q = [i_f - i_o + k_1\left(\frac{di_f}{dt} - \frac{di_o}{dt}\right)\quad k_1 c(v_i - bi_f - v_f)\quad -k_1 cai_f\quad -k_1 ci_f] \quad (18)$$

$$\theta^{\mathrm{T}} = [\Delta c\quad \Delta a\quad \Delta b\quad \Delta a\Delta b] \quad (19)$$

$$g(x) = k_1 ca \quad (20)$$

Substituting for (14), (17), (18), and (20) in the general control law and adaptation law given in (10) and (11), the AIOFLC is obtained for MU. The block diagram of the MU proposed controller is shown in Fig. 3.

C. AIOFLC for SUs

Taking the time derivative of (5) and (6) and substituting from (1)-(4), the dynamic equation of active and reactive powers is obtained as:

$$\dot{P} = f_P + u_P + \xi_P \quad (21)$$

$$\dot{Q} = f_Q + u_Q + \xi_Q \quad (22)$$

where

$$f_P = 1.5c[(i_{f\alpha} - i_{o\alpha})i_{f\alpha} + (i_{f\beta} - i_{o\beta})i_{f\beta}] - 1.5a[(bi_{f\alpha} + v_{f\alpha})v_{f\alpha} + (bi_{f\beta} + v_{f\beta})v_{f\beta}] \quad (23)$$

$$u_P = 1.5a(v_{f\alpha}v_{i\alpha} + v_{f\beta}v_{i\beta}) \quad (24)$$

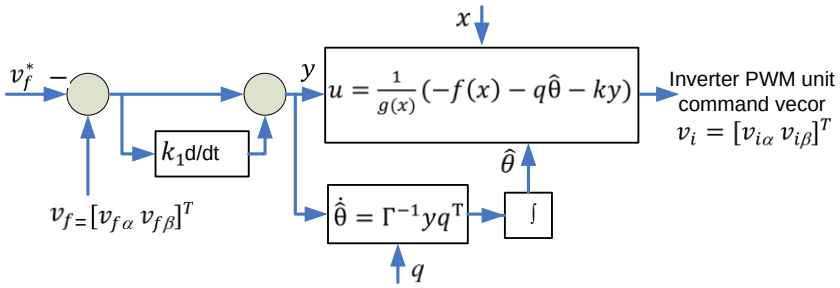


Fig. 3. Block diagram of the MU proposed controller

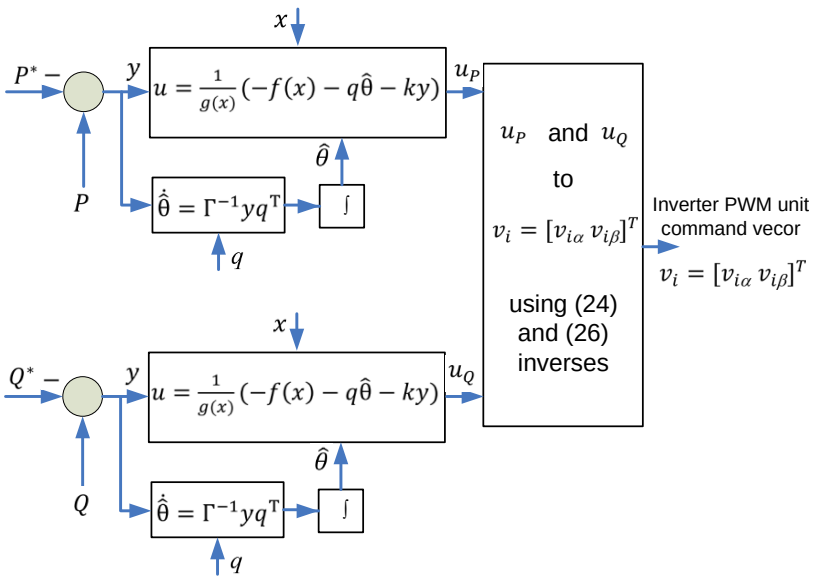


Fig. 4. Block diagram of the SUs proposed controllers

$$f_Q = 1.5c[(i_{f\beta} - i_{o\beta})i_{f\alpha} - (i_{f\alpha} - i_{o\alpha})i_{f\beta}] \quad (25)$$
$$-1.5a[(bi_{f\alpha} + v_{f\alpha})v_{f\beta} - (bi_{f\beta} + v_{f\beta})v_{f\alpha}]$$
$$u_Q = 1.5a(v_{f\beta}v_{i\alpha} - v_{f\alpha}v_{i\beta}) \quad (26)$$

ξ_P and ξ_Q are lumped uncertainties:

$$\xi_P = 1.5\xi_{v\alpha}i_{f\alpha} + 1.5\xi_{v\beta}i_{f\beta} + 1.5\xi_{i\alpha}v_{f\alpha} + 1.5\xi_{i\beta}v_{f\beta}$$
$$= \Delta c(i_{f\alpha} - i_{o\alpha})1.5i_{f\alpha} + \Delta c(i_{f\beta} - i_{o\beta})1.5i_{f\beta} + [\Delta a(v_{i\alpha} - bi_{f\alpha} - v_{f\alpha})]1.5v_{f\alpha} - (a + \Delta a)\Delta bi_{f\alpha}1.5v_{f\alpha} \quad (27)$$
$$+[\Delta a(v_{i\beta} - bi_{f\beta} - v_{f\beta})]1.5v_{f\beta} - (a + \Delta a)\Delta bi_{f\beta}1.5v_{f\beta}$$

$$\xi_Q = 1.5\xi_{v\beta}i_{f\alpha} - 1.5\xi_{v\alpha}i_{f\beta} + 1.5\xi_{i\alpha}v_{f\beta} - 1.5\xi_{i\beta}v_{f\alpha}$$
$$= \Delta c(i_{f\beta} - i_{o\beta})1.5i_{f\alpha} + \Delta c(i_{f\alpha} - i_{o\alpha})1.5i_{f\beta} + [\Delta a(v_{i\alpha} - bi_{f\alpha} - v_{f\alpha})]1.5v_{f\beta} - (a + \Delta a)\Delta bi_{f\alpha}1.5v_{f\beta} \quad (28)$$
$$+[\Delta a(v_{i\beta} - bi_{f\beta} - v_{f\beta})]1.5v_{f\alpha} - (a + \Delta a)\Delta bi_{f\beta}1.5v_{f\alpha}$$

To control the active power consider the following output

$$y = P - P^* \quad (29)$$

Taking the time derivative of (29) and using (21), gives

$$\dot{y} = f_P + u_P + \xi_P - \dot{P}^* \quad (30)$$

Considering the following functions and variables, one can achieve the canonical form (7),

$$f(x) = f_P - \dot{P}^* \quad (31)$$
$$q = [q_1\ \ q_2\ q_3\ \ q_4] \quad (32)$$
$$\theta^T = [\Delta c\ \ \Delta a\ \ \Delta b\ \ \Delta a\Delta b] \quad (33)$$
$$g(x) = 1 \quad (34)$$
$$u = u_P \quad (35)$$

where

$$q_1 = 1.5(i_{f\alpha} - i_{o\alpha})i_{f\alpha} + 1.5(i_{f\beta} - i_{o\beta})i_{f\beta} \quad (36)$$
$$q_2 = 1.5(v_{i\alpha} - bi_{f\alpha} - v_{f\alpha})v_{f\alpha} + 1.5(v_{i\beta} - bi_{f\beta} - v_{f\beta})v_{f\beta} \quad (37)$$
$$q_3 = -1.5ai_{f\alpha}v_{i\alpha} - 1.5ai_{f\beta}v_{f\beta} \quad (38)$$
$$q_4 = -1.5i_{f\alpha}v_{i\alpha} - 1.5i_{f\beta}v_{f\beta} \quad (39)$$

Now, the control and adaptation laws (10) and (11) can be used.

To control the reactive power, consider the following output

$$y = Q - Q^* \quad (40)$$

Taking the time derivative of (40) and using (22), gives

$$\dot{y} = f_Q + u_Q + \xi_Q - \dot{Q}^* \quad (41)$$

Considering the following functions and variables, one can achieve the canonical form (7),

$$f(x) = f_Q - \dot{Q}^* \quad (42)$$
$$q = [q_1\ \ q_2\ q_3\ \ q_4] \quad (43)$$
$$\theta^T = [\Delta c\ \ \Delta a\ \ \Delta b\ \ \Delta a\Delta b] \quad (44)$$
$$g(x) = 1 \quad (45)$$
$$u = u_Q \quad (46)$$

where

$$q_1 = 1.5(i_{f\beta} - i_{o\beta})i_{f\alpha} - 1.5(i_{f\alpha} - i_{o\alpha})i_{f\beta} \quad (47)$$

$$q_2 = 1.5(v_{i\alpha} - bi_{f\alpha} - v_{f\alpha})v_{f\beta} - 1.5(v_{i\beta} - bi_{f\beta} - v_{f\beta})v_{f\alpha} \quad (48)$$

$$q_3 = -1.5ai_{f\alpha}v_{i\beta} + 1.5ai_{f\beta}v_{f\alpha} \quad (49)$$

$$q_4 = -1.5i_{f\alpha}v_{i\beta} + 1.5i_{f\beta}v_{f\alpha} \quad (50)$$

Now, the control and adaptation laws (10) and (11) can be used. The block diagram of the SUs proposed controllers is shown in Fig. 4.

IV. Case studies and operation scenarios

Simulation studies are performed in MATLAB SIMULINK to investigate the designed control scheme for the MG given in Fig. 1. The parameters used in simulations are presented in Table 1. The SIMULINK solver is configured as a fixed-step with a discrete sample time of 5μs. The MU rated power and the inverters rated voltages are used as base values. The operation of the MG is evaluated using the designed controls in several cases.

A. Case Study 1

This case study is the normal testing operation for the proposed controls. In this case study, from t=0s to t=0.15s, only the MU delivers the demand powers. From t=0.15s to t=0.25s, the active power of the SU increases from 0 to 0.4pu linearly. From t=0.15s to t=0.2s, the reactive power of the SU increases from 0 to 0.1pu linearly. Fig. 5 presents the voltages and powers of three buses (i.e. PC1, PC2, and Load3) of this case study. The obtained results show the active/reactive powers track their reference values correctly. It is seen that the MU power decreases by increasing the SU power. The voltages are sinusoidal with very low harmonics and the MU bus voltage is controlled properly. No voltage deviation is observed and the responses are acceptable without any low-frequency oscillations. The voltage THDs are below the 2.5% required by IEEE 1547 and IEC 61727 standards (50% of the current harmonic limits) [24].

B. Case Study 2

In this case, all of the settings are similar to the ones given in case study 1; except that, the controllers are FLCs. Fig. 6 illustrates the obtained results of this case study. Unlike case study 1, some large oscillations are seen on active/reactive powers of MU and Load3 which are the consequences of the noticeable increases and decreases in voltages. In fact, using FLC, the MU voltage tracking is not as good as the tracking in case study 1. Besides, FLC is not a robust control against disturbances and uncertainties.

C. Case Study 3

In this case study, again all of the settings are similar to the ones considered in case 1, but the applied controls are modified CSMCs. The simulation results of this case study are shown in Fig. 7. The chattering effect and high control efforts are very noticeable. The active/reactive powers are affected really and contain high-frequency contents. VSIs filters cannot eliminate these high-frequency contents completely. The chattering phenomena can also excite unmolded dynamics which may cause instability in the real physical system. Besides, it increases the power loss of VSI switches.

TABLE 1
MG PARAMETERS

The nominal (base) power	3 MVA
VSIs nominal voltage	600 V
MG frequency	50 Hz
Line1	0.35+j0.785 Ω
Line2	0.25+j0.625 Ω
Line3	0.1+j0 Ω
VSIs DC voltage	1500 V
Switching frequency	2 kHz
Filter resistance	0.002 Ω
Filter inductance	500 μH
Filter capacitance	400 μF

The control signals of the three controllers are also shown in Figs. 3, 4, and 5. These waveforms illustrate the lower control effort of the AIOFLC.

D. Case Study 4

Nonlinear and switching loads such as rectifiers increase harmonic distortions in the power system. This case study is fully identical to case study 1, except that a rectifier, which is a nonlinear load, is paralleled to Load3 at time 0.35s. Fig. 8 shows the obtained results of this case study. Using the proposed AIOFLCs reveals that despite the harmonics in Line3 current, the voltages remain sinusoidal, and MU and SU provide the demand powers without any considerable glitch. Using the AIOFLCs, the MG has a smooth and acceptable operation. The effect of harmonic load on voltages and powers is insignificant due to the proposed high-performance controls. This case study also illustrates the robustness of the designed AIOFLCs.

E. Case Study 5

This case is fully similar to case 1; however, very higher impedances are used for Line1 and Line2: (0.35+j0.785)*6 Ω and (0.25+j0.625)*6 Ω, respectively. Fig. 9 illustrates the obtained results of this case study. The simulation results show that the MG maintains its stability; however, the harmonic contents of bus voltages are very high and the active/reactive powers are very fluctuating. This test represents that the MS strategy is only acceptable for islanded MGs in which the impedances of the lines are not very high. It can also be used in power electronics systems with paralleled VSIs.

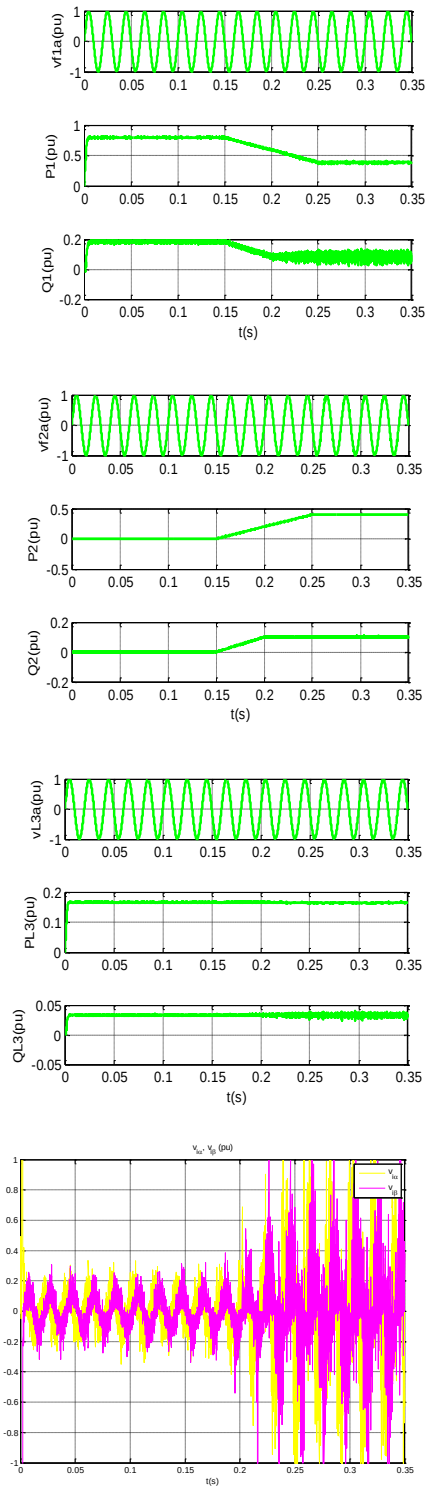


Fig. 5. AIOFLC: voltage, active/reactive powers of buses, and control signals of DG1

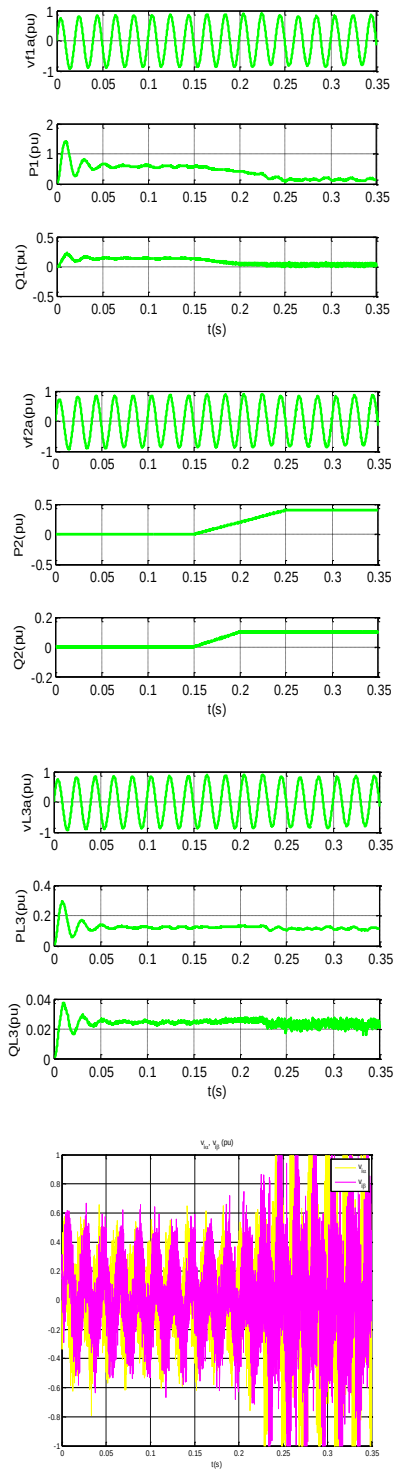


Fig. 6. FLC: voltage, active/reactive powers of buses, and control signals of DG1

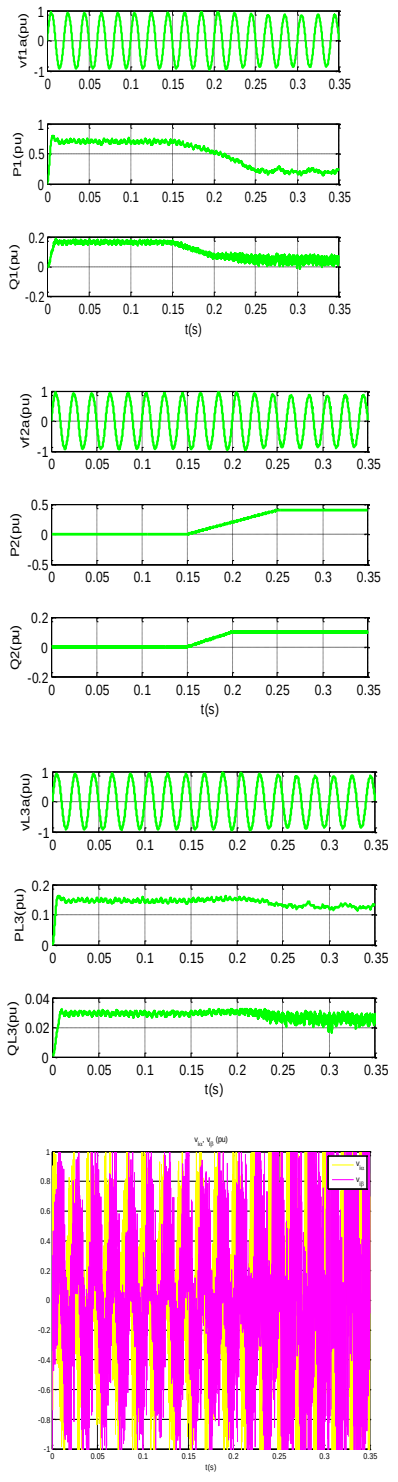


Fig. 7. Modified CSMC: voltage, active/reactive powers of buses, and control signals of DG1

F. Case Study 6

In this case study, using the parameters of case studies 1 to 3, the load switching is studied. Load2 on bus2 is disconnected at 0.3 second and connected again at 0.4 second. The simulation results of the proposed AIOFLC are represented in Fig. 10 and compared to the ones of two other standard nonlinear control methods. It is seen that using AIOFLC, the active/reactive powers of DG2 are controlled with fewer ripples, and the active/reactive powers of DG1 are also properly changed according to the load variations.

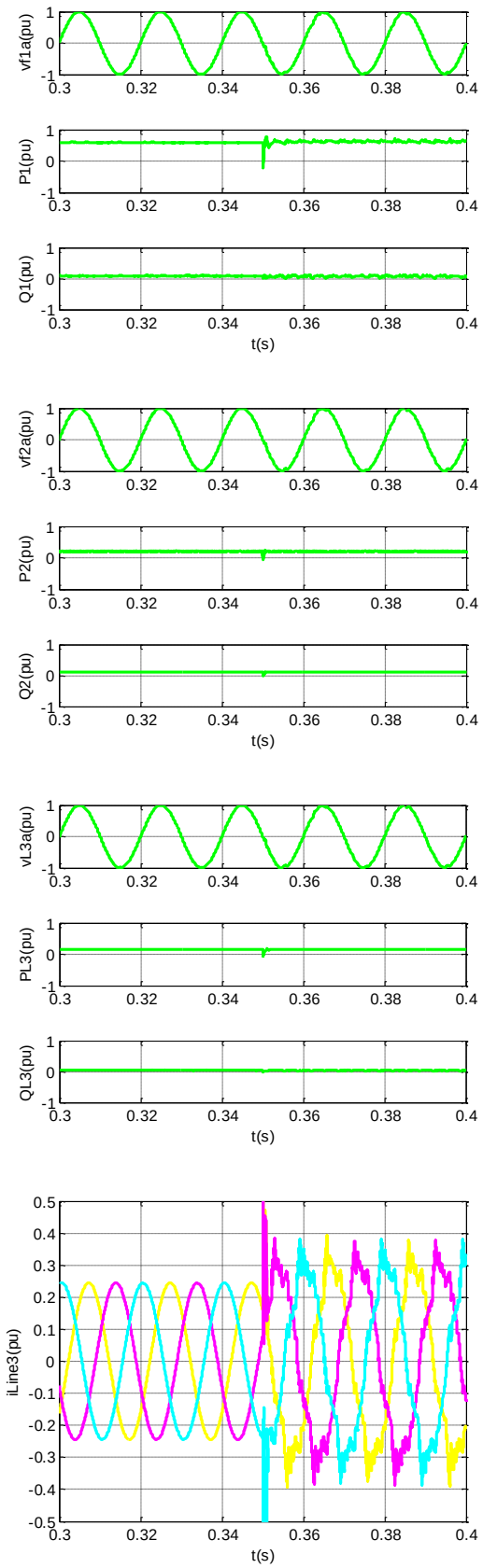


Fig. 8. AIOFLC with harmonic load (rectifier): voltage, active/reactive powers of buses, currents of the Line3

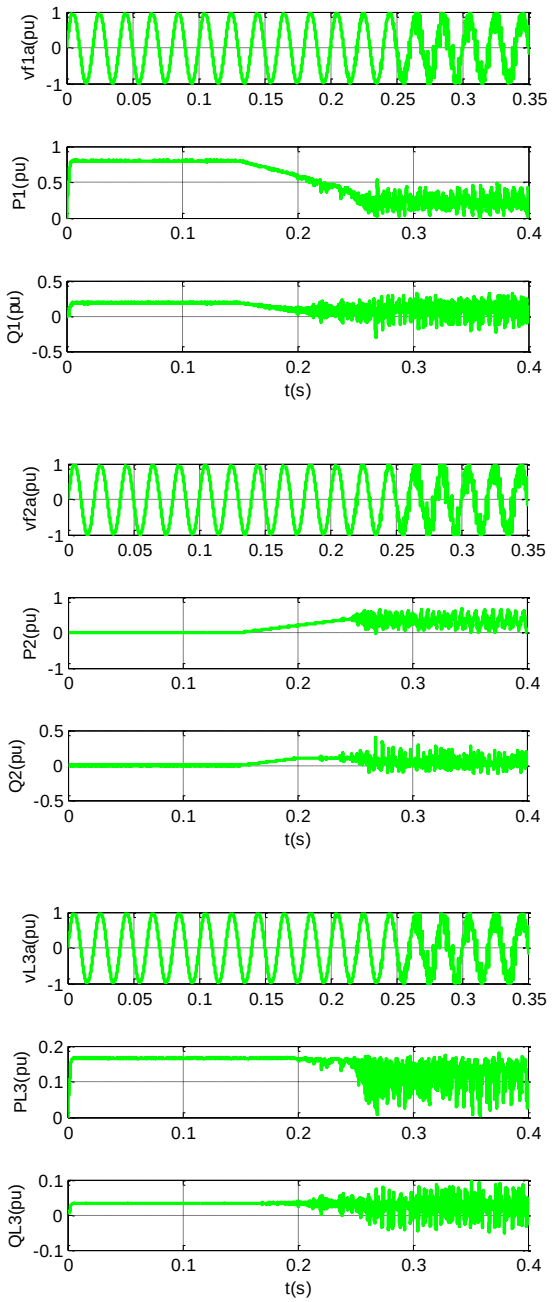


Fig. 9. High impedance lines: voltage, active/reactive powers of buses

V. Conclusions

This paper investigates AIOFLC to control islanded inverter-based MGs with MS strategy. Two AIOFLCs are designed for MU and SUs. Simulation results are obtained for an islanded MG with two DGs; however, the proposed controls can be used well for any islanded inverter-based MG with any number of DGs. Next, the responses of three controls (i.e., AIOFLC, FLC, and a modified CSMC) are obtained and compared. The obtained results illustrate the benefits of the designed

AIOFLCs. The MU bus voltage and the delivered powers of the SU are controlled properly.

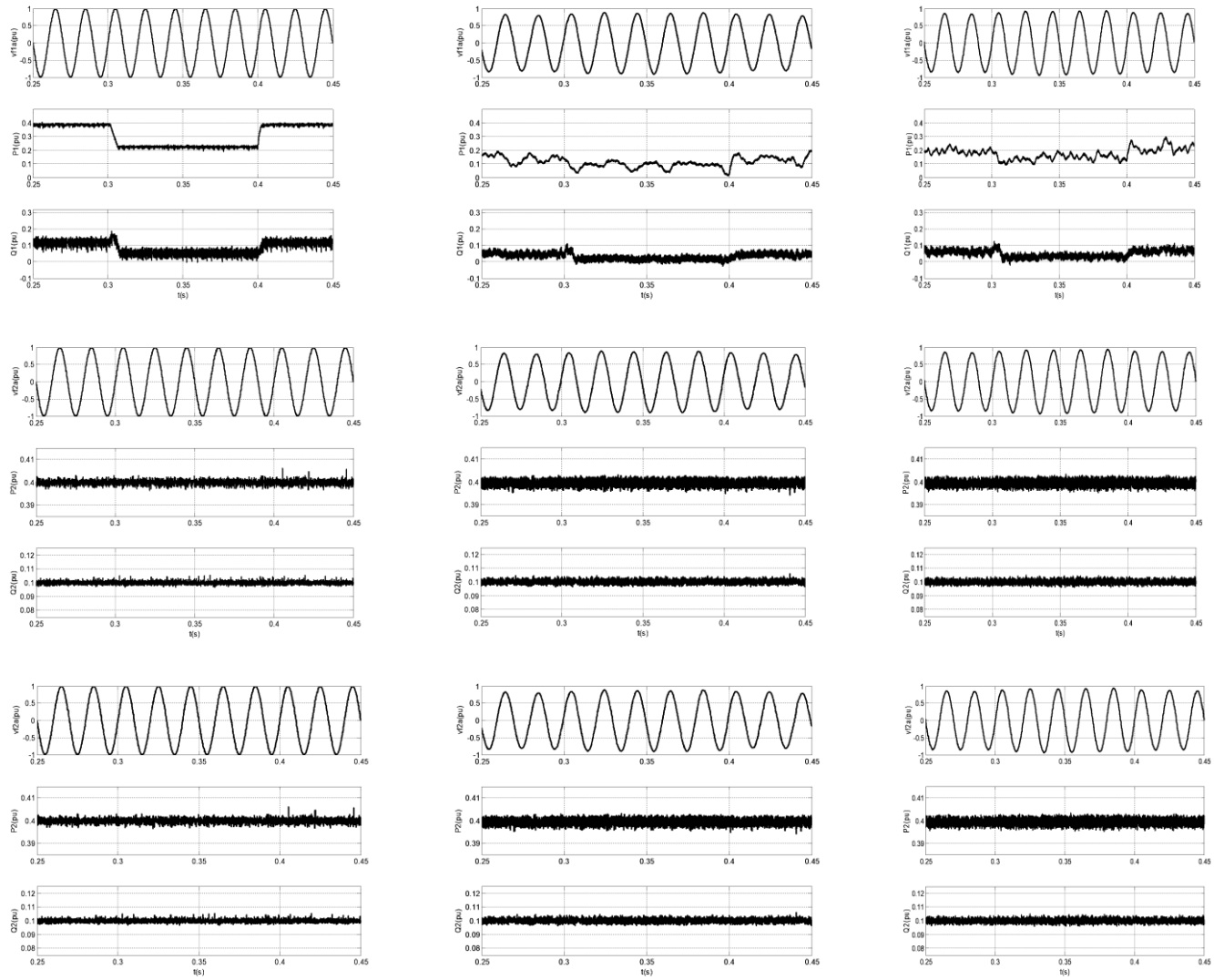


Fig. 10. Load switching: voltage, active/reactive powers of buses 1&2, (a) AIOFL, (b) FLC, (c) CSMC

By changing the loads or the active/reactive powers references of the SU, the MU generates the remaining load demand and a proper load sharing between VSIs is achieved. Using the proposed control scheme leads to almost no overvoltage or undervoltage during the transient intervals. The designed controllers provide a high-performance control behavior. In addition, simulating the volatility inherent to the DGs illustrates that the DGs can inject their powers when available. Overall, the merits of the proposed control scheme are as follows: perfect tracking, avoiding high control effort, and robustness of the AIOFLCs. The designed AIOFLCs can be used in parallel inverters with LC filters to control the voltage and power. The contributions of the paper can be summarized as follows:

-
- Rearranging the equations of inverter-based DGs with LC filters, suitable for controller design
- Designing two AIOFLCs for the MU and the SUs
- Investigating the proposed controls in a MG by adopting the MS strategy

REFERENCES

[1] A. Alfergani, A. Khalil, “Modeling and Control of Master-Slave Microgrid with Communication Delay,” The 8 th International Renewable Energy Congress (IREC 2017), 2017.

[2] J. Pinto, A. Carvalho, V. Morais, “Power Sharing in Island Microgrids,” Frontiers in Energy Research, Vol. 8, Article 609218, pp. 1-14, 2021.

[3] P. J. dos Santos Neto, T. A.S. Barros, J. P.C. Silveira, E. R. Filho, J. C. Vasquez, J. M. Guerrero, "Power management techniques for grid-connected DC microgrids: A comparative evaluation," Applied Energy,269, 115057, pp. 1-15, 2020.

[4] M. Hamzeh, H. Karimi, H. Mokhtari, "A New Control Strategy for a Multi-Bus MV Microgrid Under Unbalanced Conditions," IEEE Trans. On Power Systems, Vol. 27, No. 4, pp. 2225-2232, Nov. 2012.

[5] H. Han, X. Hou, J. Yang, J. Wu, M. Su, J. M. Guerrero, "Review of Power Sharing Control Strategies for Islanding Operation of AC Microgrids," IEEE Trans. On Smart Grid, pp. 1-16, 2015.

[6] G. B. Narejo, B. Acharya, R. S. S. Singh, and F. Newagy, Microgrids Design, Challenges, and Prospects, Taylor & Francis Group, LLC, 2022.

[7] T. Caldognetto, P. Tenti, "Microgrids Operation Based on Master-Slave Cooperative Control," IEEE Journal of Emerging and Selected Topics in Power Electronics, Vol. 2, No. 4, pp. 1081 – 1088, 2014.

[8] M. S. Mahmoud , O. Al-Buraiki , "Two-level Control for Improving the Performance of MicroGrid in Islanded Mode," 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE), Istanbul, Turkey, pp. 54-59, 2014.

[9] P. Monshizadeh, C. De Persis, N. Monshizadeh, A. V. D. Schaft, "A Communication-Free Master-Slave Microgrid with Power Sharing," 2016 American Control Conference (ACC), Boston, USA, 2016.

[10] H. Liang, Y. Dong, Y. Huang, C. Zheng, P. Li, "Modeling of Multiple Master–Slave Control under Island Microgrid and Stability Analysis Based on Control Parameter Configuration," Energies, Vol. 11, No. 9, pp. 1-18, 2018.

[11] J. Marchgraber, W. Gawlik, "Investigation of Black-Starting and Islanding Capabilities of a Battery Energy Storage System Supplying a Microgrid Consisting of Wind Turbines, Impedance- and Motor-Loads," Energies, Vol. 13, No. 19, pp. 1-24, 2020.

[12] T. Yao, R. Ayyanar , "Variable Structure Robust Voltage Regulator Design for Microgrid Master-Slave Control," 2017 IEEE Energy Conversion Congress and Exposition (ECCE), Cincinnati, OH, USA, 2017.

[13] M. Babazadeh and Houshang Karimi, "Robust Decentralized Control for Islanded Operation of a Microgrid," 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 2011.

[14] M. Dehghani**,**T. Niknam**,** M. Ghiasi**,** H. Baghaee**,** F. Blaabjerg**,** T. Dragicevĭć and M. Rashidi, "Adaptive backstepping control for master-slave AC microgrid in smart island," Energy, Vol. 246, 2022.

[15] M. M. Rezaei, J. Soltani, "Robust control of an islanded multi-bus microgrid based on input – output feedback linearisation and sliding mode control," IET Generation, Transmission & Distribution, Vol. 9, No. 15, pp. 2447–2454, 2015.

[16] M. Cucuzzella, G. P. Incremona, A. Ferrara, "Design of Robust Higher Order Sliding Mode Control for Microgrids," IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Vol. 5, No. 3, pp. 393-401, 2015.

[17] J. Wang, .Z. Liu, J. Liu, T. Wu, "A Mode Switching-Based Decentralized Secondary Control for Microgrids With Hybrid Droop and Master-Slave Structure," IEEE Open Journal of Power Electronics, Vol. 3, pp. 334-347, 2022.

[18] F. Carnielutti, M. Aly, M. Norambuena, J. Rodriguez, "Model Predictive Control for Master-Slave Inverters in Microgrids," IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society, Brussels, Belgium, 2022.

[19] S. Shajari, R. Keypour, "A New Enhanced Droop Controller for Seamless Load Sharing in AC Microgrids in Presence of Wind Turbine and Photovoltaic Sources," International Journal of Industrial Electronics, Control and Optimization, Vol. 5, No. 2, pp. 153-165, April 2022.

[20] M. Alizadeh, H. Askarian, A. Bakhshai, N. Khodabakhshi-Javinani, "An Enhanced Distributed State Feedback for Secondary Control in an Islanded Microgrid," International Journal of Industrial Electronics, Control and Optimization, Vol. 5, No. 2, pp. 123-132, April 2022.

[21] N. R. Abjadi, "Nonsingular Terminal Sliding Mode Control for Islanded Inverter-Based Microgrids," Journal of Operation and Automation in Power Engineering, Vol. 12, No. 1, pp. 26-34, Jan. 2024.

[22] S. A. A. Fallahzadeh, N. R. Abjadi, A. Kargar, "Decoupled Active and Reactive Power Control of a Grid-Connected Inverter-Based DG Using Adaptive Input–Output Feedback Linearization," Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020.

[23] Z. Ding, Nonlinear and Adaptive Control Systems, Published by The Institution of Engineering and Technology, London, United Kingdom, 2013.

[24] R. Teodorescu, M. Liserre, P. Rodriguez, Grid Converters for Photovoltaic and Wind Power Systems, John Wiley & Sons, Ltd, United Kingdom, 2011.

Navid Reza Abjadi received the B.S., M.S., and Ph.D. degrees in electrical engineering from Isfahan University of Technology, Isfahan, Iran in 1999, 2001, and 2010 respectively. He is a lecturer at Shahrekord University. His research interests include motor drives, control theory applications, and power electronics.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Active Harmonic Compensation and Stability Improvement in High Power Grid-Connected Inverters Using Unified Power Quality Conditioner

Elham Samavati[1] | Hamid Reza Mohammadi[2,*]

Faculty of Electrical and Computer Engineering, University of Kashan, Kashan, Iran.[1,2]
* Corresponding author's email: mohammadi@kashanu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 7—Jun-2023
Received in revised form: 7-Sep-2023
Accepted: 8-Sep-2023
Published online: 18-Sep-2023

Keywords:
United Power Quality Conditioner (UPQC),
Grid-Connected Inverter,
Weak Grid,
Impedance-Based Stability,
Nonlinear Load.

ABSTRACT

Objective: Multifunctional features of grid-connected inverters can be used for harmonic compensation of local load voltage and grid-injected current. But, in high-power grid-connected inverters, there is a challenge due to low switching frequency. On the other hand, simultaneous compensation of local load voltage and grid-injected current harmonics is an important issue in grid-connected inverters. Using a Unified Power Quality Conditioner (UPQC) at the Point of Common Coupling (PPC), an improved active harmonic compensation method is proposed which is appropriate for high-power low-frequency grid-connected inverters. The UPQC operates as a combination of a negative shunt virtual admittance and a negative series virtual impedance at the PCC. It suppresses the disturbances caused by local load variation and grid impedance change. Using a low-power, high-frequency UPQC, local load voltage and grid-injected current harmonics up to higher-order components are simultaneously compensated despite grid impedance changes and nonlinear local load variations. The control system is designed according to the impedance-based stability criterion to ensure the system's stability. The theoretical results are validated using different case study simulations in MATLAB/Simulink software.

Nomenclature

Y_T	Equivalent admittance at the PCC	i_{o_h}	Harmonic components of the inverter output current
Y_o	Output admittance of the grid-connected inverter	U_{PCC_h}	Harmonic components of the PCC voltage
Y_l	Variable admittance of the local load	i_I	The output current of the inverter
Z_g	The grid impedance	I_s	The current source in the grid-connected inverter model
U_g	The grid voltage		
U_{PCC}	PCC voltage		
U_l	Nonlinear local load voltage		
I_g	The grid-injected current		

I. Introduction

Along with the widespread utilization of renewable energy sources in power systems, large changes have occurred at transmission and distribution levels. Nowadays, the harmonic components generated by nonlinear loads and inverter-based distributed generation (DG) units, highly threaten the quality of grid-injected current.

A. Research Motivation

To alleviate problems caused by variations in nonlinear local load and grid impedance, appropriate power quality compensators should be designed [1]. The grid-connected inverter in inverter-based DG units should be designed and controlled in such a way that the grid-injected current harmonics are minimized. The IEEE 1547 standard determines the permissible limits for harmonic contents of the PCC voltage and grid-injected current [2]. If the THD of the PCC voltage increases over its permissible limit, then the THD of the grid-injected current increases, and consequently, the protective devices operate and the DG unit disconnects from the grid. The main challenge is related to the effect of changes in nonlinear local load and grid impedance on the increasing harmonic components and harmonic instability of the DG unit. On the other hand, the grid impedance variation causes resonant frequency variation. In this condition, the grid-connected inverter is stable if the ratio of the equivalent grid impedance to the inverter output impedance satisfies the Nyquist stability criterion [3].

B. Literature Review

Several control methods were recently proposed for suppressing harmonic disturbances at the PCC. Some of these methods are such as the grid voltage feedforward methods [4-5], H^{∞} repetitive controller [6-7], using auxiliary inverters [8-10], virtual impedance-based control techniques [11-21], and methods that shape the inverter output impedance [22-25]. Also, to solve the resonant problem, the capacitor current feedback was used in an active damping control loop to improve the system stability [26-27].

In [8], the authors used an auxiliary inverter containing a series LC filter. The parameter adjustment of the series LC filter increases the main inverter output impedance and consequently suppresses the harmonic components at the PCC. In [9] two shunt inverters were used to compensate for the grid-injected current and the PCC voltage harmonics, simultaneously. The local load voltage harmonics were reduced by the first inverter and the current harmonics caused by the nonlinear local load were reduced by the second one. Meanwhile, a parallel virtual admittance can be applied to the output of the inverter to compensate for the grid-injected current harmonics. Also, a virtual impedance can be applied in series to eliminate the voltage harmonics at the PCC [10-14]. Harmonic reference voltages were generated according to the harmonic virtual inductance and resistance in negative and positive sequences [15-17]. In reference [18], using an adaptive virtual impedance, the harmonic components of the load voltage were suppressed. The load voltage distortion was used to determine this virtual impedance. Another strategy for harmonic suppression is inverter output impedance shaping. If the equivalent output impedance of a current-controlled grid-connected inverter is set to infinite, the grid-injected current harmonics are eliminated. On the other hand, if the equivalent series impedance at the PCC is set to zero, the grid voltage harmonics do not affect the grid-injected current [19-23]. To reduce both the PCC voltage and grid-injected current harmonics, a hybrid virtual impedance method was proposed in [24]. In [25], the application of a series active power filter was introduced to compensate for the PCC voltage harmonics in case of grid impedance change. However, the nonlinear local load was not considered in this work. In previous works, the harmonic compensation method was usually designed based on the unchanged virtual admittance. However, the nonlinear local load and the grid impedance may be practically varied for several reasons. These variations cause variable current and voltage harmonics at the PCC. So, the value of applied virtual admittance should be adaptively changed.

This problem was investigated in our previous work where a single-phase low-power grid-connected inverter was considered [28]. In this reference, a variable virtual admittance was suggested, and it was implemented using the multifunctional capability of the grid-connected inverter. The switching frequency of the grid-connected inverter was high, and therefore the high-frequency components could be compensated. Also, in this reference, a series active power filter was used to compensate for the grid impedance effect. But, in high-power three-phase grid-connected inverters, the switching frequency is low, and therefore they cannot be used for harmonic compensation at high-frequencies.

In [31] a new adaptive harmonic voltage control method was developed for voltage-controlled DG inverters, which neither uses any PI regulators nor imposes stability issues associated with nonideal implementation of infinite gains of PR controllers. The developed control logic can be used for both grid-connected and islanding modes and is capable of accurate output voltage tracking by exploiting the property of an optimal switching vector controller.

In [32], a new strategy was designed based on multiple resonant current controllers and active damping feedback to improve the ability of the PV inverter to suppress harmonics under the grid background harmonic conditions. A passive impedance network was constructed using the impedance model of a PV inverter and then by analyzing the series resonance of the impedance network, the amplification coefficient of the harmonic voltage at the point of common coupling (PCC) was obtained. To solve the problem that the output harmonics exceed the standard limits under the background harmonic condition of the weak grid, a harmonic mitigation control strategy was implemented.

An improved fuzzy-controlled back-to-back electric spring was proposed in [33] that was used to improve power quality in microgrids. The proposed structure was controlled to operate simultaneously as an electric spring and shunt active power filter. That means, the series part of the back-to-back electric spring regulates the critical load voltage and applies the demand-side management and the parallel part operates as a shunt active power filter capable of power factor correction and current harmonic compensation.

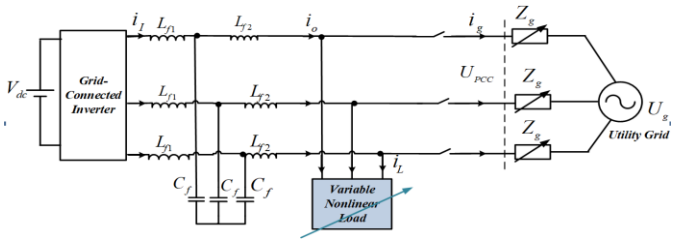


Fig. 1. Grid-connected inverter system with a variable nonlinear load that connected to a weak grid

C. The Necessity of the Research

In real conditions, the grid impedance and nonlinear local load may be changed. So, it is needed to propose a method that improves the power quality and stability in this condition. Also, in high-power grid-connected inverters, because of low switching frequency, the grid-connected inverter cannot be used for implementation of the virtual admittance concept to improve the power quality and stability. Therefore, this paper aims to propose a new UPQC-based structure for grid-connected inverters to compensate for the voltage and current harmonics and improve the system's stability in this condition.

D. Novelty and Main Contribution

In this paper, a low-frequency three-phase grid-connected inverter with an LCL filter has been considered under local load variation and weak grid conditions as shown in Fig. 1.

To implement a variable virtual admittance in parallel with the inverter and a variable virtual impedance in series with the grid, the application of a low-power high-frequency UPQC at the PCC is proposed. The proposed UPQC-based structure is capable of compensating the grid-injected current and PCC voltage harmonics in case of local load and grid impedance variations. The UPQC is the combination of a shunt and a series inverter which are connected in the DC-link. The shunt inverter is modeled as a harmonic current source and the series inverter is modeled as a harmonic voltage source. The series inverter injects a series voltage to compensate for the variable grid impedance effect and the shunt inverter injects a shunt current to compensate for the variable nonlinear local load effect. With the proposed UPQC-based structure, it is possible to suppress both the PCC voltage and grid-injected current harmonics and improve the system's stability.

Using the proposed structure in a low-frequency grid-connected inverter, the robustness of the system against the nonlinear local load and grid-impedance variations is maintained. The results of the proposed method were compared with the results without using the proposed structure.

E. Organization and Structure

The rest of the paper is organized as follows. The principle of the proposed harmonic cancellation method is described in section 2. In section 3, the control block diagram of the proposed system is introduced and then the simulation results are presented in section 4. Finally, concluding remarks are given in section 5.

II. Principle of the Proposed Harmonic Cancellation Method

The equivalent circuit of a grid-connected inverter is shown in Fig. 2. The grid-connected inverter is modeled as a current source (I_s) in parallel with the inverter output admittance (Y_o). In this model, I_g is the grid-injected current, U_l is the nonlinear local load voltage which is equal to the PCC voltage (U_{PCC}) and finally, U_g is the grid voltage. It is assumed that the nonlinear local load is variable. The grid impedance (Z_g) includes the transformer leakage impedance, the line impedance, and the output impedance of other inverters connected to the PCC. The grid impedance is variable because the grid structure is changed and the number of grid-connected inverters may be varied. If the local load is modeled as a variable admittance (Y_l), then equation (1) expresses the equivalent admittance at the PCC.

$$Y_T = Y_o + Y_l \quad (1)$$

According to Fig. 2, U_l and I_g can be written as (2) and (3), respectively. These equations show that the local load and grid impedance changes cause variable harmonics at the PCC. To solve this problem, the effects of the local load and grid impedance variations should be canceled.

$$U_l = I_s\left(\frac{Z_g * \frac{1}{Y_T}}{\frac{1}{Y_T} + Z_g}\right) + U_g\left(\frac{\frac{1}{Y_T}}{\frac{1}{Y_T} + Z_g}\right) \quad (2)$$

1.1.

$$I_g = I_s\left(\frac{\frac{1}{Y_T}}{\frac{1}{Y_T} + Z_g}\right) - U_g\left(\frac{1}{\frac{1}{Y_T} + Z_g}\right) \quad (3)$$

A. Cancel Out the Grid Impedance Effect

An LCL filter has a resonant frequency that is varied due to the grid impedance variation. So, the harmonic stability analysis is important in weak grids. To overcome this challenge, an effective active damping method is necessary to guarantee the system's stability when connected to a weak grid. As shown in Fig. 3, the effect of grid impedance can be canceled out by applying $-Z_g$ as a series virtual impedance. This is done using a low-power series active power filter. According to (2) and (3), using the proposed cancellation method, U_l and I_g can be written as (4) and (5), respectively.

$$\lim_{Z_g \to 0} U_l = U_g \quad (4)$$

$$\lim_{Z_g \to 0} I_g = I_s - U_g Y_T \quad (5)$$

Equation (4) shows that the grid impedance effect on local load voltage harmonics is eliminated. Also, equation (5) shows that the grid impedance effect on grid-injected current harmonics is well eliminated. However, the local load variation effect on harmonic components of the grid-injected current is not canceled out due to Y_T presence.

B. Nonlinear Local Load Effect Cancellation

As shown in Fig. 2, the equivalent admittance at the PCC (Y_T) includes the parallel connection of inverter output admittance and variable local load admittance. To cancel out the effect of local load variation, an adaptive virtual admittance concept can be realized

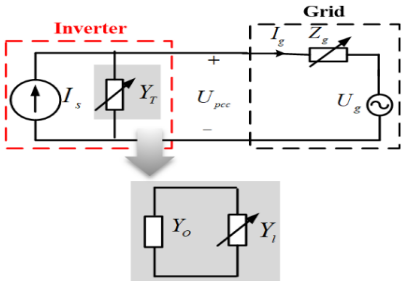


Fig. 2. Equivalent model of an inverter with a variable local load connected to a weak grid

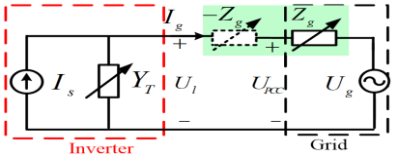


Fig. 3. Grid impedance cancellation method using series virtual impedance concept

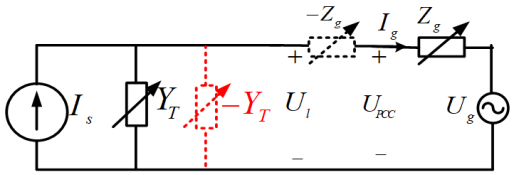


Fig. 4. Equivalent circuit model of a grid-connected inverter uses variable virtual admittance and variable virtual impedance concepts

using a shunt active power filter, as shown in Fig. 4. In this figure, the equivalent circuit model of a grid-connected inverter is shown which uses the proposed virtual admittance and virtual impedance concepts. Using the proposed active cancellation method, the undesired effects of nonlinear local load and grid impedance are removed. Therefore, the local load voltage (U_l) and grid-injected current (I_g) can be written as follows:

$$\lim_{\substack{Z_g\to 0\\ Y_T\to 0}} U_l = U_g \tag{6}$$

$$\lim_{\substack{Z_g\to 0\\ Y_T\to 0}} I_g = I_s \tag{7}$$

Regarding (6) and (7), by elimination of the nonlinear local load and grid impedance effects, the local load voltage (U_l) and grid-injected current (I_g) are exactly similar to the grid voltage and current reference of the DG unit, respectively. Following the grid-injected current harmonic compensation, the PCC voltage (U_{PCC}) harmonic components are reduced.

The harmonic stability of the proposed structure can be analyzed using the impedance-based stability criterion [29]. The bandwidth of the Phase Lock Loop (PLL) is usually less than 50 Hz, which is

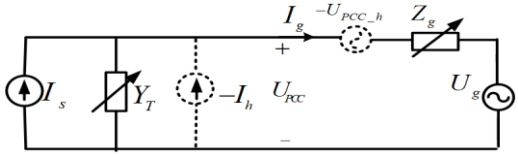


Fig. 5. Equivalent circuit model of a grid-connected inverter with the proposed UPQC-based structure

much less than the bandwidth of the internal current control loop (usually more than 1 kHz). So, the PLL effect is not considered in the output impedance of the grid-connected inverter. To check the impedance-based stability criterion, equation (3) is rewritten as follows:

$$I_g=(I_s - Y_T U_g) * \left(\frac{1}{1+\frac{Z_g}{\frac{1}{Y_T}}}\right) \tag{8}$$

According to equation (8), a grid-connected inverter will be stable if the ratio of the grid impedance to the inverter output impedance including nonlinear local load ($Z_g \ /(\frac{1}{Y_T})$ or $Z_g * Y_T$), satisfies the Nyquist criterion. Because of grid impedance cancellation using a series active power filter and nonlinear local load cancellation using a shunt active power filter, $Z_g * Y_T$ becomes zero and satisfies the Nyquist criterion. So, the grid-connected inverter will be stable in all conditions.

C. *Principle Of UPQC*

The active cancellation method which is proposed in previous sections is implemented using a low-power, high-frequency UPQC at the PCC. As shown in Fig. 5, the UPQC is operated as a shunt current source and a series voltage source, simultaneously.

The shunt current source injects the harmonic components of the inverter output current in the opposite sign. So, the harmonic components of the grid-injected current are canceled out at the PCC. Also, the series voltage source injects the PCC voltage harmonics in the opposite sign. So, the harmonic components of the local load voltage are canceled out at the PCC.

III. Control Block Diagram of the Proposed System

An LCL-filtered grid-connected inverter that uses the proposed UPQC-based structure is shown in Fig. 6. The grid impedance and the nonlinear local load are assumed to be variable. A PR controller is used to control the grid-connected inverter. The grid-injected current (i_g), the output current of the inverter (i_l) and the PCC voltage (U_{PCC}) are inputs of the controller. The control block diagram of the grid-connected inverter is shown in Fig. 7.

The *dq* reference currents are assumed to be i_d=100 and i_q=0. A PR controller with a transfer function of $G_{PR\_f}(s) = k_p + \sum_{i=1,5,7} \frac{2k_i\omega_c s}{s^2+2\omega_c s+\omega_0^2}$ is designed to minimize the steady-state error of the grid-injected current (i_g) in *αβ* reference frame. The pole-zero cancellation method is used to design the controller parameters. The shunt and series inverters in the UPQC structure are used to produce the harmonic components of the inverter output current (i_{o_h}) and PCC voltage (U_{PCC_h}), respectively. Low-power, high-frequency inverters in the UPQC structure are capable of harmonic compensation up to higher orders. The control block diagram of the series inverter is shown in Fig. 8. To eliminate the grid impedance variation effect, the harmonic content of the PCC voltage (U_{PCC}) is extracted using the *dq* reference frame and is inversely injected into the system using the series inverter. $G_{d2}(s)$

is the delay transfer function. Also, the control block diagram of the shunt inverter is

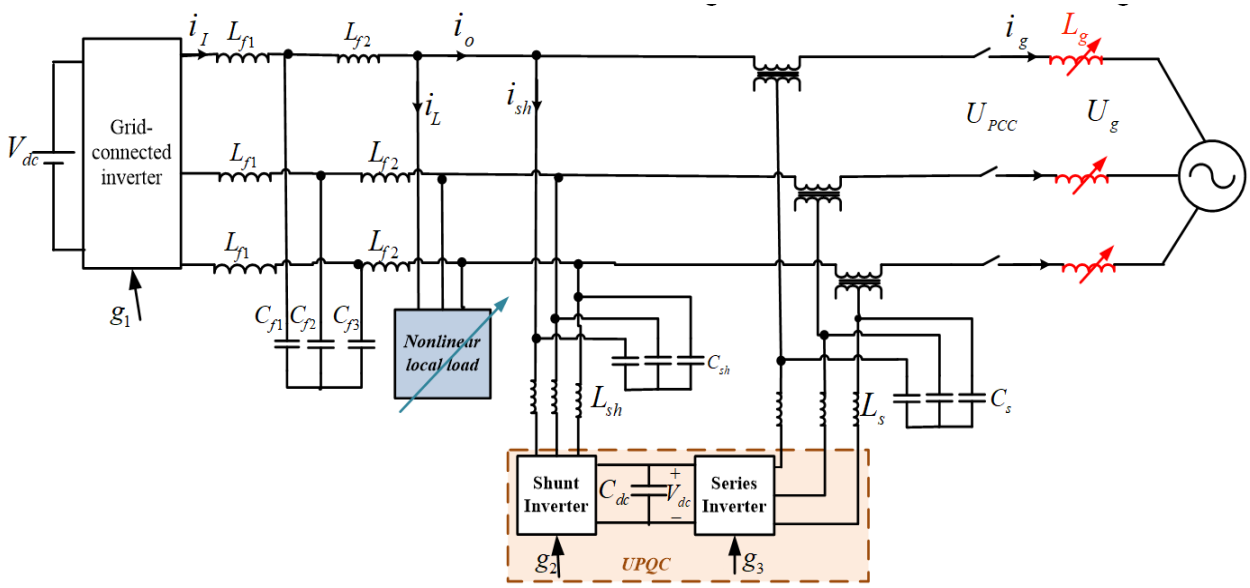


Fig. 6. An LCL-filtered grid-connected inverter that uses the proposed UPQC-based structure

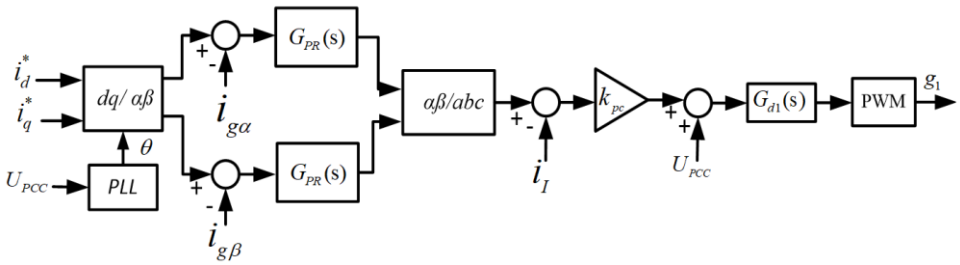


Fig. 7. Control block diagram of the grid-connected inverter

shown in Fig. 9. This control system is designed to eliminate the current harmonics and regulate the DC-link voltage. The grid-injected current harmonics are extracted using the dq reference frame and are used to generate the reference current. Therefore, the shunt inverter compensates for the harmonic components of the grid-injected current at the PCC. To regulate the DC-link voltage, a proportional-integral (PI) controller is used in a voltage feedback control loop. The output of the PI controller is added to the d component of the reference current. The shunt inverter is controlled using the hysteresis current control method. High-frequency switching harmonics in the outputs of the shunt and series inverters are damped using LC filters.

IV. Simulation Results

In this section, a 380V/50Hz three-phase grid-connected inverter that uses the proposed UPQC-based structure, as shown in Fig. 6, is simulated using MATLAB/Simulink software. Also, the control block diagram of the main grid-connected inverter is modeled based on Fig. 7, and the control block diagram of the series and shunt inverters in the UPQC structure are modeled based on Fig. 8 and Fig. 9. The simulation parameters are given in Table 1.

The grid inductance is changed between 0 and 7.2mH and a variable local load is connected to the inverter output. It is validated that the proposed structure has an excellent performance in case of local load and grid impedance changes.

The harmonic compensation capability of the proposed structure is evaluated in case the grid inductance is equal to 7.2mH and the nonlinear load2 (shown in Fig. 10(b)) is connected. Without using the UPQC-based structure, the PCC voltage and the grid-injected current waveforms and their THDs are shown in Fig. 11(a). As shown in this figure, the voltage and current THDs have exceeded the standard limits.

Also, the simulation results in the case of using the proposed UPQC-based structure are shown in Fig. 11(b). As shown in this figure, the voltage and current THDs are satisfied by the IEEE 1547 standard under the worst conditions.

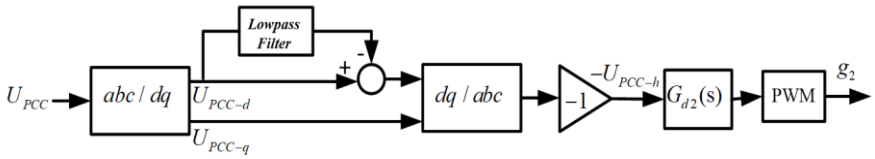


Fig. 8. Control block diagram of the series inverter

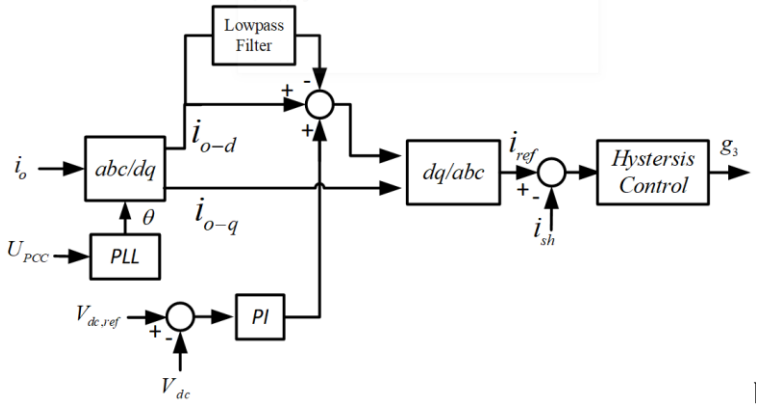


Fig. 9. Control block diagram of the shunt inverter

TABLE 1
SIMULATION PARAMETERS

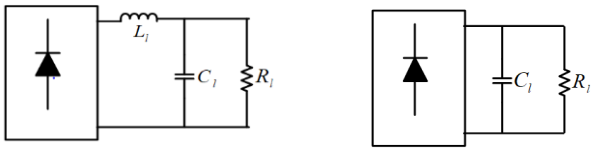
Symbol	Parameter	Value	Symbol	Parameter	Value
V_{dc1}	DC-link voltage of the grid-connected inverter	750V	L_g	Grid inductance	0∼7.2mH
L_{f1}	Inverter side inductance	5.6mH	V_{dc2}	DC-link voltage of the UPQC	90V
L_{f2}	Grid side inductance	2μH	L_l	Nonlinear load inductance	300mH
C_f	Filter capacitance	45μF	R_l	Nonlinear load resistance	20Ω
L_{sh}	LC filter inductance of the shunt inverter	2mH	C_l	Nonlinear load capacitance	300μF
C_{sh}	LC filter capacitance of the shunt inverter	0.5μF	K_p	Proportional gain of the PR controller	0.2
L_s	LC filter inductance of the series inverter	3.5mH	K_i	Harmonic frequency resonant gains of the PR controller	9,5,4
C_s	LC filter capacitance of the series inverter	0.5μF	ω_c	The cut-off frequency of the PR controller	4Hz
C_{dc}	DC-link capacitance of the UPQC	4×2000μF	K_{pc}	Controller gain	2.5
U_{in}	The line-to-line voltage at PCC	380V	f_s	Switching frequency	2500Hz

A. *Local Load Variation*

At First, a three-phase linear resistive load (R=55Ω) is connected to the output of the inverter. At t=0.1s, the resistive load is replaced with a nonlinear load (load1 as shown in Fig. 10(a)). Also, at t=0.2s, load1 is replaced with another nonlinear load (load2 as shown in Fig. 10(b)). The load current is shown in Fig. 12. The PCC voltage (U_{PCC}) and the grid-injected current (i_g) without and with using the proposed UPQC structure are shown in Figs. 13 and 14, respectively.

Without using the proposed UPQC structure, the THD of the PCC voltage and THD of the grid-injected current have exceeded the permissible limits.

However, as shown in Fig.14, when using the proposed UPQC structure, the THD values and harmonic factors of the PCC voltage



(a) (b)

Fig. 10. Nonlinear loads (a) load1 (b) load2

1.2. *and grid-injected current are acceptable according to the limits determined by the IEEE 1547 standard.*

B. *Grid Impedance Variation*

1.3. In this subsection, the grid impedance variation effect on the performance of a grid-connected inverter with a nonlinear local load (load2) is investigated. Firstly, the grid impedance is equal to 4mH. At t=0.1s, it is increased to 7.2mH, and at t=0.2s it is decreased to 5mH. Waveforms and frequency spectrums of U_{PCC} *and* i_g, without and with using the *proposed UPQC structure* are shown in Figs. 15

and 16, respectively. As shown in Fig. 15 the THD of the grid-injected current is decreased when the grid inductance is increased. This is because the grid inductance acts as a first-order lowpass filter. As the grid inductance increases, the crossover frequency of this filter decreases, and consequently, the THD of the grid-injected current decreases. It is observed that without using the proposed UPQC structure, THDs of U_{PCC} *and* i_g and the harmonic factors of the individual voltage and current components are beyond the permissible limits. But, when the proposed UPQC structure is used, the THD values and the harmonic factors of the individual voltage and current components are satisfied by the IEEE 1547 standard.

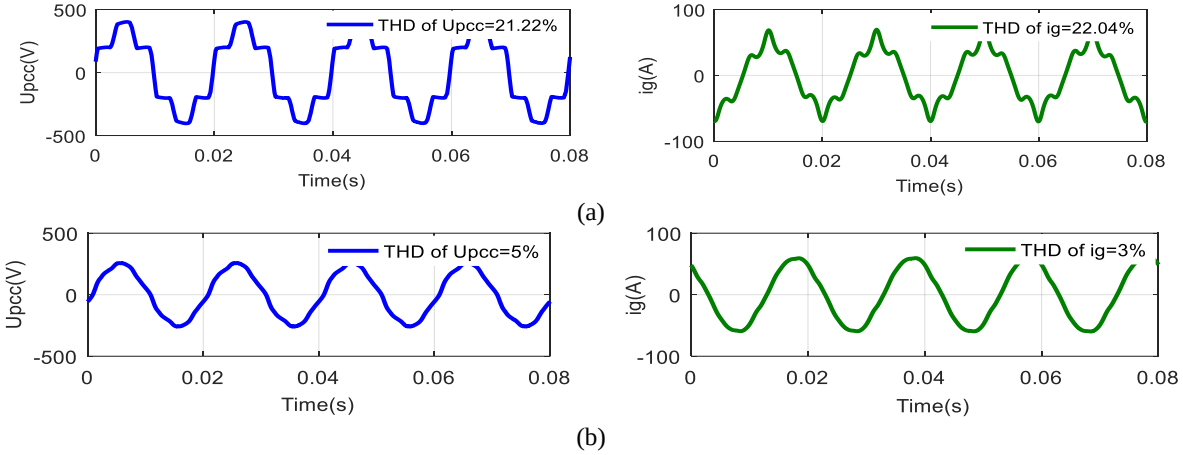


Fig. 11. PCC voltage (U_{PCC}) and grid-injected current (i_g) (a) without UPQC-based structure (b) with UPQC-based structure

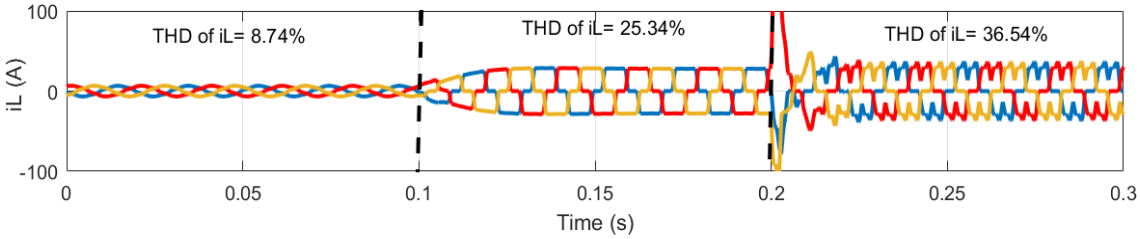


Fig. 12. Local load current in case of load variation

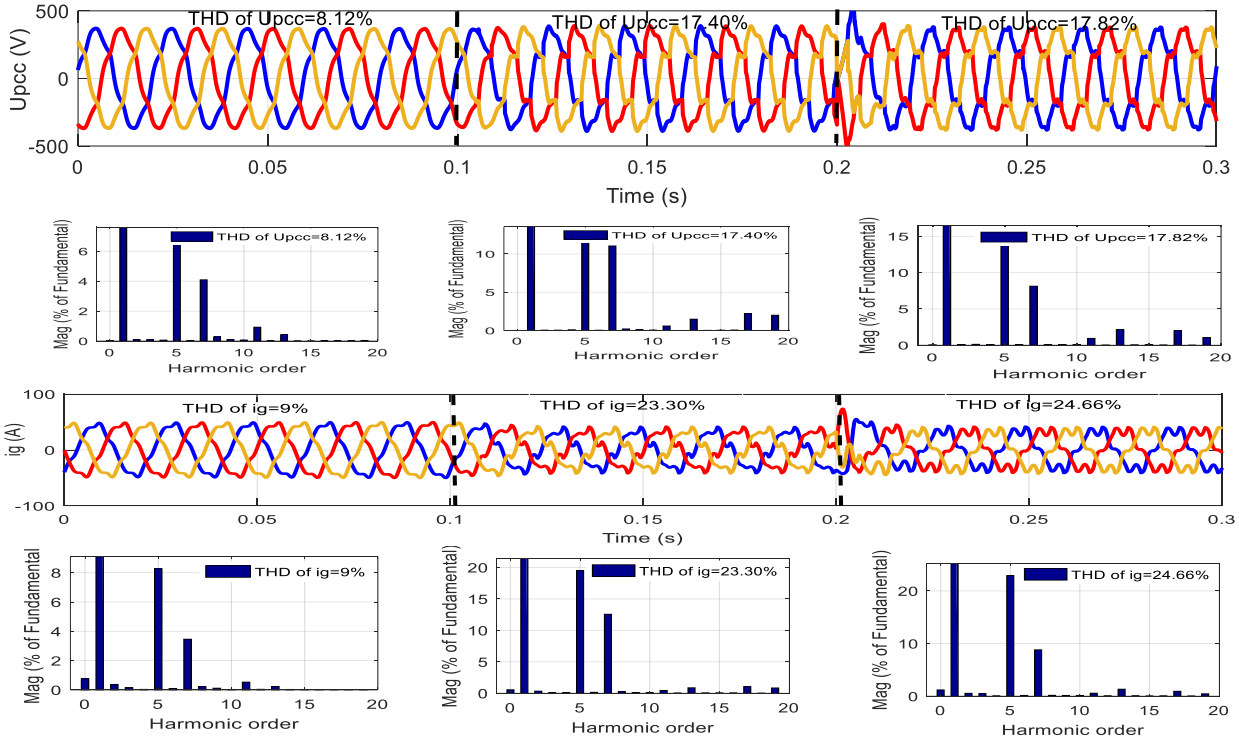


Fig. 13 PCC voltage ($\mathbf{U_{pcc}}$) and grid-injected current ($\mathbf{i_g}$) in case of variable local load without UPQC-based structure

C. Performance analysis considering the grid voltage distortion

The grid voltage distortion deteriorates the performance of grid-connected inverters. According to the IEEE 519 standard [30], the maximum permissible limit of the grid voltage THD is 8%. In this study, the THD of the distorted grid voltage is considered 7%. According to previous works, the grid voltage harmonic rejection can be done using the grid voltage or the capacitor voltage of the

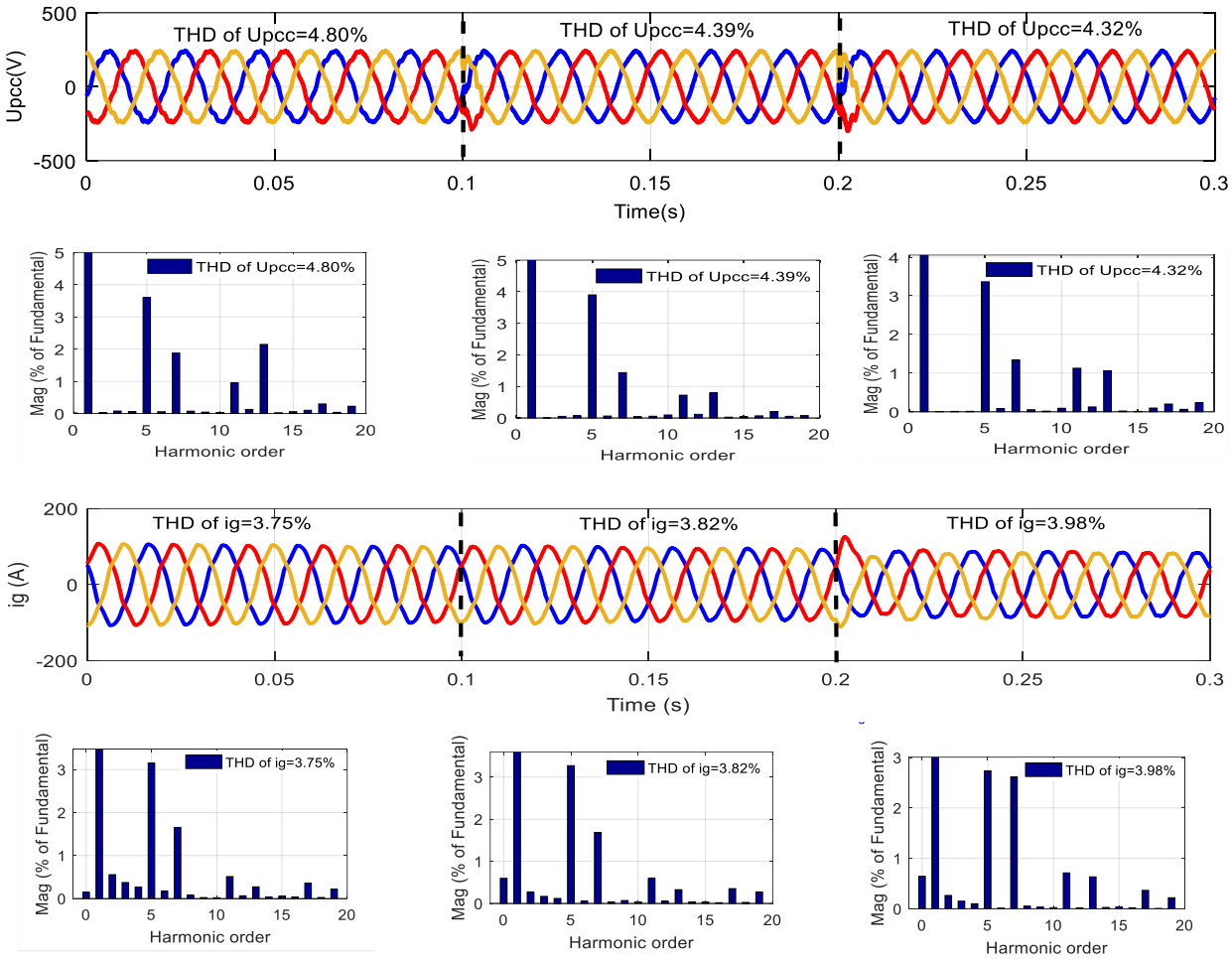


Fig. 14 PCC voltage (U_{pcc}) and grid-injected current (i_g) in case of variable local load with *UPQC-based structure*

LCL filter as a feedforward signal. As shown in Fig. 7, in the proposed strategy, the grid voltage is proportionally injected as a feedforward signal to the output of the control loop. In this case, the grid impedance is equal to 7.2mH and the nonlinear load (load2) is connected. Waveforms and frequency spectrums of the grid-injected current and the PCC voltage are shown in Fig. 17.

Regarding frequency spectrums in Fig. 17, the harmonic factors of the individual harmonic components are within permissible limits determined by the IEEE 1547 standard. The grid-injected current THD is less than 5%. Also, the proposed structure is capable of reducing the THD of the PCC voltage significantly. For better comparison, Table 2 gives the local load voltage and grid-injected current THDs for all case studies.

V. Conclusion

In this work, the variable virtual admittance and variable virtual impedance concepts were applied using a UPQC-based structure for low-frequency grid-connected inverters. The proposed structure has the capability of harmonic compensation and stability improvement in case of local load changes and variable grid impedance. The principle of this method is based on the impedance-based stability criterion for grid-connected inverters. The nonlinear local load changes and the grid impedance variations lead to the variable harmonic content of the local load voltage, PCC voltage, and grid-injected current. In high-power grid-connected inverters, the variable virtual admittance concept cannot be applied by the inverter because of the low switching frequency. Therefore, the harmonic distortion at the PCC does not satisfy the requirements of the IEEE 1547 standard. In this paper, the topic of cancellation of the negative effects caused by local load changes and grid impedance variations in a high-power three-phase grid-connected inverter was presented.

It was verified that by using a low-power high-frequency UPQC, the output admittance of the inverter and the grid impedance became zero at the point of load connection. Therefore, the harmonic components were removed even under the change in local load and grid impedance variation. Also, in the proposed method, the system stability was ensured based on the impedance-based stability criterion. The effectiveness of the proposed method was verified using different simulation results in MATLAB/Simulink software. In this work, the system under study was a balanced 3-phase 3-wire system. The unbalanced 3-phase 4-wire systems can be considered in future works.

TABLE 2
THDs OF THE LOCAL LOAD VOLTAGE AND GRID-INJECTED CURRENT (%)

		Without the proposed UPQC-based structure		With the proposed UPQC-based structure	
		local load voltage	grid-injected current	local load voltage	grid-injected current
Variable local load	Resistive load	8.12	9	4.80	3.75
	Nonlinear load	17.4	23.3	4.39	3.82
	Highly Nonlinear load	17.82	24.66	4.32	3.98
Variable grid impedance	Z_g=4mH	18.61	29.63	4.37	3.81
	Z_g=7.2mH	19.05	20.87	2.64	3.86
	Z_g=5mH	18.99	25.22	2.85	4.14
Considering grid voltage distortion		-	-	6.44	4.31

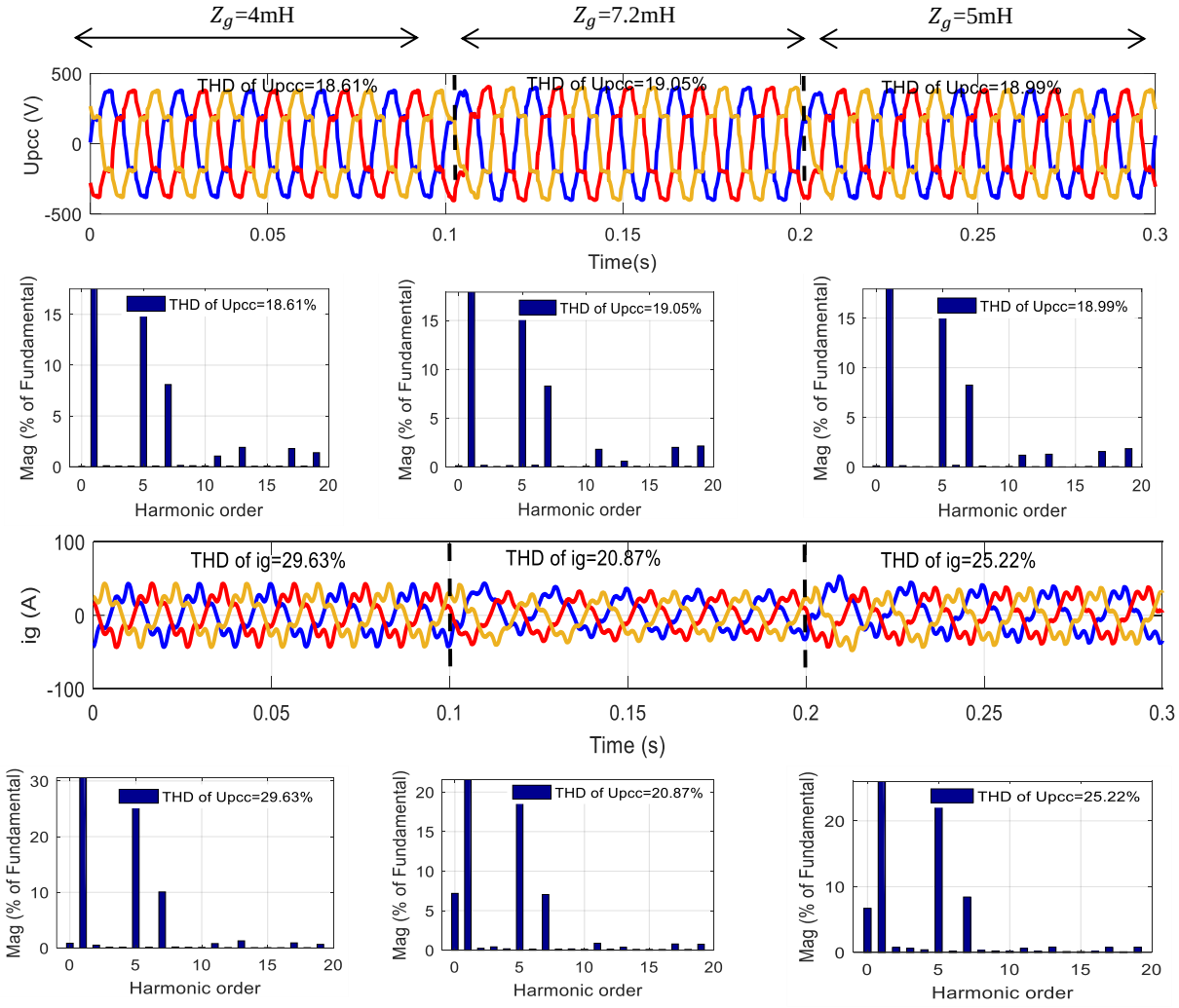


Fig. 15 PCC voltage ($\mathbf{U_{pcc}}$) and grid-injected current ($\mathbf{i_g}$) in case of variable grid impedance without UPQC-based structure

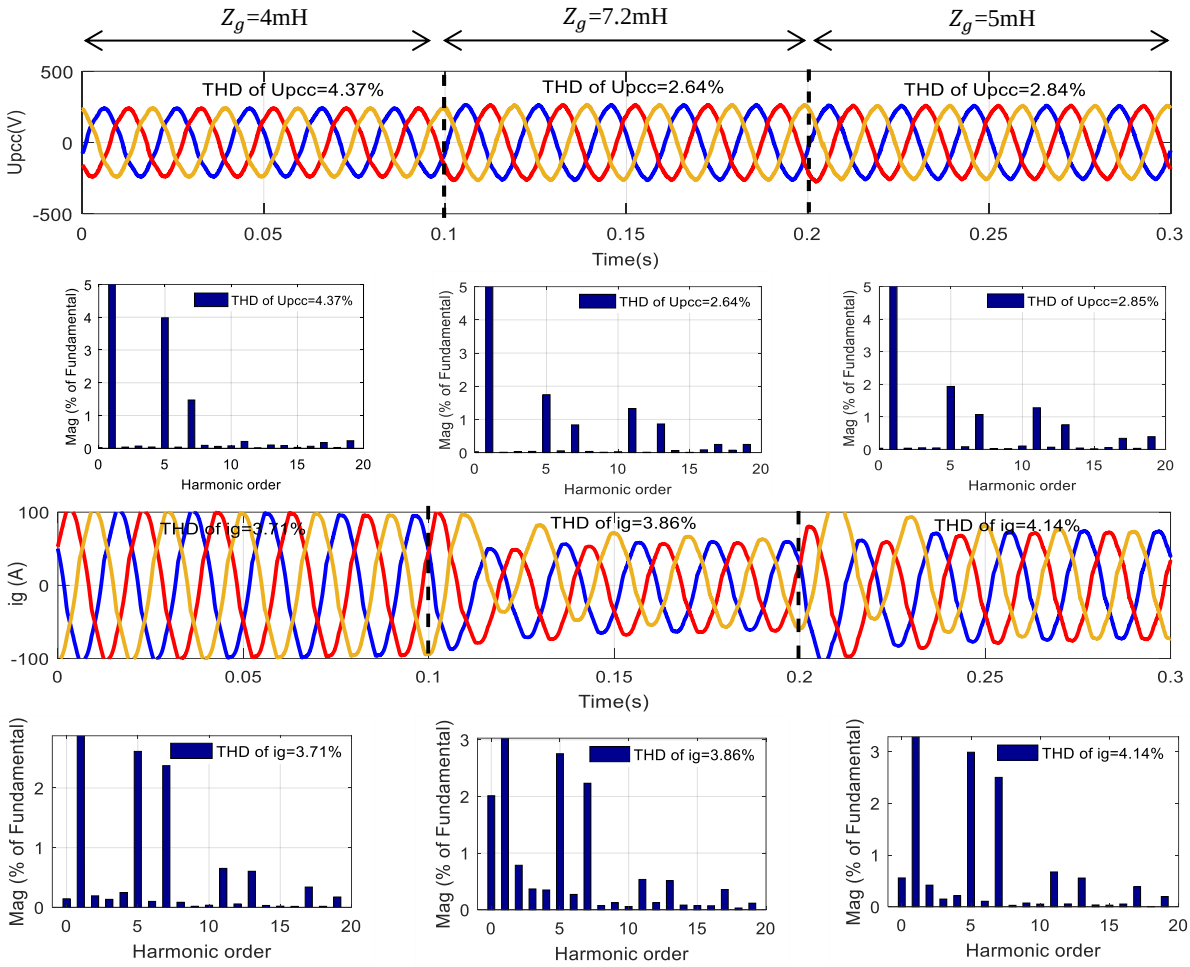


Fig. 16 PCC voltage (U_{PCC}) and grid-injected current (i_g) in case of variable grid impedance with UPQC-based structure

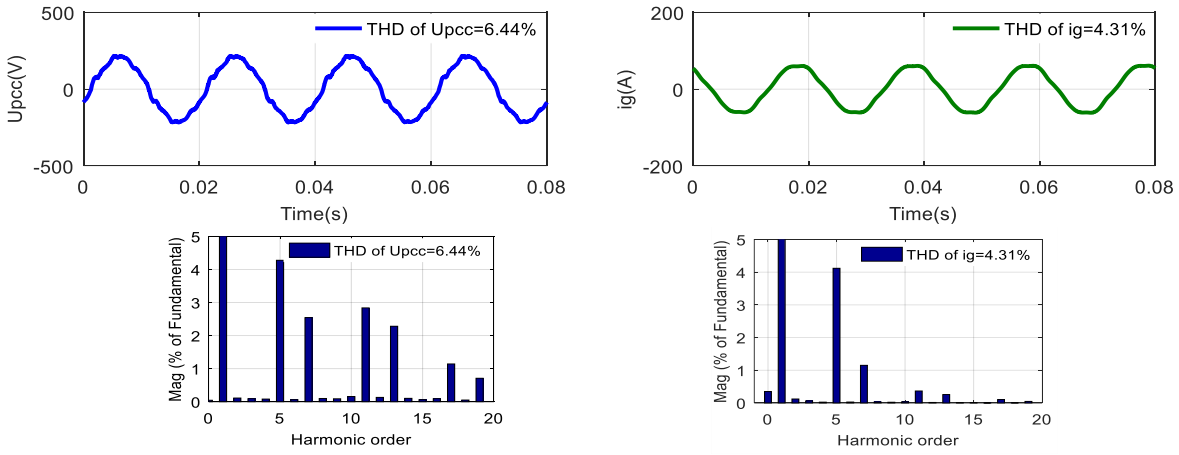


Fig. 17 PCC voltage (**U_{pcc}**) and grid-injected current (**i_g**) considering the grid voltage distortion

REFERENCES

[1] Sun K., Wang N., Wei Li, Y., Nejabatkhah F., Mei Y., Lu X., 'Parallel Operation of Bi-directional Interfacing Converters in a Hybrid AC/DC Microgrid under Unbalanced Grid Voltage Conditions', IEEE Transactions on Power Electronics, 2016, 32, (3), pp. 1872 – 1884.

[2] IEEE Std. 1547-2003: 'IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems', 2003.

[3] Enslin J., Heskes P., 'Harmonic interaction between a large number of distributed power inverters and the distribution grid', IEEE Transactions Power Electronics, 2004, 19, (6), pp. 1586-1593.

[4] He Y., Wang K., Chung H., 'Utilization of Proportional Filter Capacitor Voltage Feedforward to Realize Active Damping for Digitally-Controlled Grid-Tied Inverter Operating under Wide Grid Impedance Variation', in Proc. IEEE ECCE, 2014, pp. 4450-445

[5] Wang X., Ruan X., Liu S., Tse C. K., 'Full Feedforward of Grid Voltage for Grid-Connected Inverter with LCL Filter to Suppress Current Distortion Due to Grid Voltage Harmonics', IEEE Transactions Power Electronics, 2010, 25, (12), pp. 3119-3127.

[6] Zhong Q. C., Hornik T., 'Cascaded Current –Voltage Control to Improve the Power Quality for a Grid-Connected Inverter with a Local Load', Transactions on Industrial Electronics, 2013, 60, (4), pp. 1344-1355.

[7] Trivedi A., Singh M., 'Repetitive Controller for VSIs in Droop Based AC-Microgrid', IEEE Transactions on Power Electronics, 2017, 32, 8, pp. 6595-6604.

[8] Bai H., Wang X., Blaabjerg F., Loh P. C., 'Harmonic Analysis and Mitigation of Low-Frequency Switching Voltage Source Inverter with Auxiliary VSI', IEEE Journal of Emerging and Selected Topics in Power Electronics, 2018, 6, 3, pp. 1355-1365.

[9] He J., Liang B., Li Y.W., Wang C., 'Simultaneous microgrid voltage and current compensation using coordinated control of dual-interfacing-converters', IEEE Transactions on Power Electronics, 2017, 32, (4), pp. 2647-2660.

[10] Javadi A., Woodward L., Al-Haddad K.l., 'Real-time Implementation of a Three-phase THSeAF Based on a VSC and a P+R Controller to Improve Power Quality of Weak Distribution Systems', IEEE Transactions on Power Electronics, 2018, 33, (3), pp. 2073-2082.

[11] Ebrahimzadeh E., Blaabjerg F., Wang X., Bak C. L., 'Bus Participation Factor Analysis for Harmonic Instability in Power Electronics Based Power Systems', IEEE Transactions on Power Electronics, 2018, 33, (12), pp. 10341-10351.

[12] Blanco C., Reigosa D., Vasquez J. C., Guerrero J. M., Briz F., 'Virtual Admittance Loop for Voltage Harmonic Compensation in Microgrids', IEEE Transactions on Industry Applications, 2016, 52, (4), pp. 3348-3356.

[13] He J., Li Y., Blaabjerg F., Wang X., 'Active Harmonic Filtering Using Current-Controlled, Grid-Connected DG Units with Closed-Loop Power Control', IEEE Transactions on Power Electronics, 2014, 29, (2), pp. 642-653.

[14] Li Y. W., He J., 'Distribution System Harmonic Compensation Methods: An Overview of DG-Interfacing Inverters', IEEE Industrial Electronics Magazine, 8, (4), pp. 18-31.

[15] Savaghebi M., Guerrero J. M., Jalilian A., Vasquez J. C., 'Mitigation of Voltage and Current Harmonics in Grid-Connected Microgrids', 2012 IEEE International Symposium on Industrial Electronics **Conference,** Hangzhou, China**,** 2012.

[16] Wang X., Blaabjerg F., Chen Z., 'Synthesis of Variable Harmonic Impedance in Inverter-Interfaced Distributed Generation Unit for Harmonic Damping Throughout a Distribution Grid', 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), Orlando, FL, USA, 2012, pp. 775-782.

[17] Savaghebi M., Shafiee Q., Vasquez J. C., Guerrero J. M., 'Adaptive Virtual Impedance Scheme for Selective Compensation of Voltage Unbalance and Harmonics in Microgrids', 2015 IEEE Power & Energy Society General Meeting, Denver, CO, USA, July 2015.

[18] Micallef A., Apap M., Spiteri-Staines C., Guerrero J. M., 'Mitigation of Harmonics in Grid-Connected and Islanded Microgrids Via Virtual Admittances and Impedances', IEEE Transactions on Smart Grid, 2017, 8, (2), pp. 651-661.

[19] Ruan X., 'Control Techniques for LCL-Type Grid-Connected Inverters', Springer Nature Singapore Pte Ltd. and Science Press, 2018, pp. 227-248.

[20] Rold´an-P´erez J., Rodr´ıguez-Caberoy A., Prodanovi´cz M., 'Harmonic Virtual Impedance Design for Parallel-Connected Grid-Tied Synchronverters', IEEE Journal of Emerging and Selected Topics in Power Electronics, 2019, 7, (1), pp. 493-503.

[21] Gadelovits S., Zhong Q. C., Kadirkamanathan V., 'Cascaded Control to Shape Output Virtual Impedance and Improve Output Voltage Quality for Power Inverter', 2016 IEEE 7th International Symposium on Power Electronics for Distributed Generation Systems (PEDG) **Conference**, Vancouver, BC, Canada, 2016, pp. 1-6.

[22] Gadelovits S., Zhong Q. C., Kadirkamanathan V., Kuperman A., 'UDE-based Controller Equipped with a Multi-Band-Stop Filter to Improve the Voltage Quality of Inverters', IEEE Transactions on Industrial Electronics, 2017, 64, (9), pp. 7433-7443.

[23] Ni R., Li Y. W., Zhang Y., Zargari N. R., Cheng, Z., 'Virtual Impedance-Based Selective Harmonic Compensation (VI-SHC) PWM for Current Source Rectifiers', IEEE Transactions on Power Electronics, 2014, 29, (7), pp. 3346-3356.

[24] Wang T., Nian H., Zhu Z. Q., Ding L., Zhou B., 'Flexible Compensation Strategy for Voltage Source Converter Under Unbalanced and Harmonic Condition Based on a Hybrid Virtual Impedance Method', IEEE Transactions on Power Electronics, 2018, 33, (9), pp. 7656-7673.

[25] He Y., Chung C., Liang B., Zhang X., Wu W., 'Active Cancellation of Equivalent Grid Impedance for Improving Stability and Injected Power Quality of Grid-Connected Inverter under Variable Grid Condition', IEEE Transactions on Power Electronics, 2018, 33, (11), pp. 9387-9398.

[26] Zhao X., Meng L., Xie C., Guerrero J. M., Wu X., Vasquez J. C., Savaghebi M., 'A Voltage Feedback Based Harmonic Compensation Strategy for Current-Controlled Converters', IEEE Transactions on Industry Applications, 2018, 54, (3), pp. 2616-2627

[27] Lim K., Choi J., Jang J., Moon S., Kim J., 'An Improved Method for Harmonic Mitigation and Stability Improvement of Grid-Connected Inverters under Output Voltage Regulation for Harmonic Compensation under Islanded Mode of Microgrid', 2014 16th International Power Electronics and Motion Control Conference and Exposition **Conference,** Antalya, Turkey, 2014.

[28] Samavati E., Mohammadi H.R., 'An Improved Method for Harmonic Mitigation and Stability Improvement of Grid-Connected Inverters under Local Load Variation and Weak Grid Condition', International Journal of Electrical Power and Energy Systems, 123, 2020

[29] Sun J., 'Impedance-Based Stability Criterion for Grid-Connected Inverters', IEEE Transactions on Power Electronics, 2011, 26, (11), pp. 3075-3078.

[30] IEEE Std 519-2014, 'IEEE recommended practice and requirements for harmonic control in electric power systems', 2014.

[31] Srinivas V.L., Singh B., Mishra S., Xu L., 'Harmonic Voltage Control in Distributed Generation Systems Using Optimal Switching Vector Strategy', IEEE Systems Journal, 2022, 16, (2), pp. 1861 – 1872.

[32] Zhao E., Han Y., Lin X., Liu E., Yang P., Zalhaf A. S. 'Harmonic Characteristics and Control Strategies of Grid-connected Photovoltaic Inverters under Weak Grid Conditions', International Journal of Electrical Power & Energy Systems, 2022, 142, Part B.

[33] Sadeghi S., Mohammadi H.R., 'An Improved Fuzzy Controlled Back-to-Back Electric Spring Using Hybrid Structure of ES-1 and Shunt-APF to Improve Power Quality in Microgrids', International Journal of Industrial Electronics, Control and Optimization, 2022, 5, (1), pp 89-98

Elham Samavati was born in Yazd, Iran, in 1988. She received the B.S. degree from Shahid Bahonar University, Kerman, Iran, in 2011, the M.S. degree from the Khaje Nasir University, Tehran, Iran, in 2012, and the Ph.D. degree from the University of Kashan, Kashan, Iran, in 2020 all in electrical engineering. Currently, she is a researcher at Y.B.T. company. Her research interests include power electronics, power quality, system automation, renewable energies, and energy management.

Hamid Reza Mohammadi was born in Qom, Iran, in 1971. He received the B.S. degree from the Sharif University of Technology, Tehran, Iran, in 1993, the M.S. degree from the University of Tabriz, Tabriz, Iran, in 1995, and the Ph.D. degree from Tarbiat Modares University, Tehran, Iran, in 2008 all in electrical engineering. Currently, he is an Associate Professor in the Electrical Engineering Department at the University of Kashan, Kashan, Iran. His research interests include power electronics, power quality, active filters, and different microgrid aspects including modeling and control of power converters, power quality, and energy management.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Controlling the Ground Particle Size and Ball Mill Load Based on Acoustic Signal, Quantum Computation Basis, and Least Squares Regression, Case Study: Lakan Lead-Zinc Processing Plant

Sadegh Kalantari[1] | Ali Madadi[2] | Mehdi Ramezani[3] | Abdolmotaleb Hajati[4]
Department of Electrical Engineering, Tafresh University, Tafresh, Iran.[1, 2]
Department of Mathematics, Tafresh University, Tafresh, Iran.[3]
Department of Mining Engineering, Arak University of Technology, Arak, Iran.[4]
Corresponding author's email: Ramezani@tafreshu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 24-June-2023
Received in revised form: 13-September-2023
Accepted: 15-Sep-2023
Published online: 19-Sep-2023

Keywords:
Least Squares Regression,
Quantum Computation,
Ball Mill System Identification,
Acoustic Signal,
Ball Mill Control.

ABSTRACT

Grinding in a ball mill is a process with high energy consumption; therefore, a slight improvement in its performance can lead to significant economic benefits in the industry. The softness of the product of the grinding circuits prevents loss of energy in the subsequent processes. In addition, controlling the mill's performance is challenging due to its complex dynamic characteristics. The primary purpose of this article is to use the ground particle size diagram and acoustic signal in ball mill control, and model their relationship based on the least squares method. As a result, by extracting valuable data from the acoustic signal, the optimal condition of the ball mill_ in terms of ground particle size and ball mill load (standard, low, high)_ can be achieved. In doing so, in this article, innovative ideas such as adaptive quantum basis, sparse representation, SVD, and PCA-based methods were used. The proposed method has been practically implemented on the ball mill of the Lakan lead-zinc processing plant. Also, a prototype of the device was built. The test results show that the optimal load for the studied ball mill is 10t/h. In this case, the ground particle size is 110-120 microns, which is ideal for this plant. Also, the power spectrum is in the middle-frequency band (300-700 Hz). According to the analysis and results, the proposed method will increase the efficiency of the studied ball mill.

I. Introduction

A. Background

In mineral processing, particles containing valuable minerals must be ground to reach a size that is small enough to be released from the tailings and easily separated by a suitable concentration method [1, 2]. Grinding is a critical operation in mineral processing, which has a considerable impact on the economic value of the product. Many industrial surveys have proven that the grinding process accounts for a significant part of the total cost of metal production [3, 4]. Impact and friction are the primary mechanisms for size reduction in grinding. Ball milling is an essential process with more than 50% of the total energy consumption; therefore, a slight improvement in ball mill efficiency would result in a substantial economic benefit for the industry. The other important issue in the grinding process is the softness of the grinding circuit product. It affects the performance of the subsequent separation processes. Therefore, the precise control of the grinding circuits is significant to prevent energy loss and improve the efficiency

of the operations. Controlling the performance of a closed circuit ball mill is a challenging issue due to its complex dynamic characteristics and numerous interactions during the control loops [5]. Also, the variables of the ball mill cannot be easily measured due to the harsh working conditions and the impossibility of stopping it. Therefore, it is impossible to access all the variables for more accurate observation and better control of the ball mill performance. The main goal of researchers is to reduce the energy consumption of the grinding circuits, and, at the same time, to reduce the abrasion of grinding media to maintain the high quality of the final product [6, 7]. To reach this goal, a model is needed for optimal control of the ball mill regarding ground particle size and load [8]. One of the variables that has been studied by the researchers is the acoustic signal resulting from the collision of the material with the shell. The acoustic signal is a superposition of the grinding parameters, which shows the general state of the ball mill [5]. Therefore, it can indirectly express the critical condition of the ball mill operation. Although the relationship between the acoustic signal and the grinding conditions is complex, some of the ball mill characteristics can be extracted by examining the behavior of the time domain and frequency of the signal. Finally, using the acoustic signal data, the conditions inside the ball mill can be identified. Also, a suitable model can be introduced to control the ball mill system.

B. Literature Review

The acoustic signal has been studied in various researches. Recently, the relationship between sound intensity and ground particle size has been investigated in the Alumina ball mill [5]. In this paper, the critical rotation speed of the ball mill has been determined to reach the proper particle size using the sound spectrum analysis. According to this article, when the rotational speed is lower or higher than the critical speed, the size of the output particles increases, and the efficiency of the ball mill decreases accordingly. In addition, the sound intensity in critical speed is higher than the other modes. Therefore, there is a direct correlation between sound intensity and grinding efficiency [5]. Tang et al. have introduced some developments in the grinding process based on soft measuring methods such as acoustic signals and mechanical vibration [9]. In that article, techniques such as Fourier and Wavelet are reviewed to extract information from sound and vibration signals. Some offline and online models have also been presented for industrial applications. Shi et al. represented some acoustic signal measurement methods using fractional Fourier transform to detect the ball mill load [10]. The data related to load parameters are extracted by representing the signal as a fractional Fourier transform and spectrum subtractions. Characteristics of the grinding signal are obtained from the geometric spectral subtraction based on an autoregressive model in reference [11]. Using ensemble empirical mode decomposition and multi-scale spectral information, the conditions of the mill load are determined. Method [12] uses acoustic emission signals to train the convolutional neural network that can predict particle size distributions. In this method, the required data are obtained from the discrete element method and used for network training and testing.

C. Research Challenges and Motivations

Although in most of the previous research, soft measuring methods have been performed only on a laboratory sample to know the parameters of the ball mill. Here, the aim is to use such techniques in industrial applications [9]. In this paper, a new approach was implemented to control the ground particle size and feed rate (load of a ball mill) in the Lakan lead-zinc processing plant. Lakan lead-zinc processing plant is 46 kilometers southwest of Arak, a city in the central province of Iran. This factory is one of the most important sources of lead and zinc concentrate production, which started working in 1968. One of the challenges of this factory is its outdated equipment. Since the power consumption of this device is high, controlling its feed rate and output ground particle size can result in higher efficiency and lower energy consumption. Therefore, this article aims to present an effective method for optimal control of ball mills. To control the ball milling operation, parameters such as input feed rate, input water, rotational speed, and ball mill particle size distribution are the minimum requirements [8]. The specifications of the studied ball mill, whose liner is made of plastic, are shown in Table 1. It should be noted that the sizes of the balls are not fixed, ranging from 20 to 90 mm. The reason is that if a fixed size is used, some holes will be created between the balls, increasing the grinding time and the size of the ground particles. As illustrated, the rotational speed and input water of the ball mill studied in this article are constant; however, the feed rate can vary. Therefore, in this study, by changing the feed rate of the ball mill, we achieved optimal efficiency, and the feed rate was proved to be a critical parameter for optimal control of the ball mill.

D. Methodology

In this study, the wet ball mill efficiency of the Lakan lead-zinc processing plant is investigated according to the feed rate and the ground product, before and after the ball charging. To identify the ball mill system, it is necessary to sample the acoustic signal.

TABLE 1
Characteristics of the Ball Mill

Length	2500 mm
Outer diameter	2300 mm
Gearbox input speed	1775 rpm
Output speed	21 rpm
Ball size range	20-90 mm
Thickness of body and liners	1500 mm
Liner material	Plastic

Therefore, the acoustic signal is first recorded at different times (based on retention time and under different feed rates) while the mill operates before and after the ball charge. In the meantime, the output product of the ball mill is sampled, and after sieve analysis, the distribution of ground particle size is obtained. Then, the outliers of the recorded acoustic signal are removed using a method based on quantum computation and sparse representation. In the next step, the relationship between the audio signal and ground particle size distribution is obtained using least square regression, processing the sound signal and extracting its helpful features, such as the maximum frequency component. By applying this model, the load and ground particle size ball mill can be effectively controlled. In this study, after identifying the ball mill system using an acoustic signal, an electronic control device was also built and tested. According to the test results, the presented method improved the wet ball mill efficiency of the Lakan lead-zinc processing plant. The sections of the article are organized as follows. In section 2, the proposed method is entirely explained. In section 3, the results and practical tests are presented, and conclusions are given in section 4.

II. Proposed Method

In this part, all the steps of the proposed method will be explained in detail. The proposed plan includes modeling and a ball mill control scheme. The modeling steps and ball mill control flowcharts are shown in Fig. 1 and Fig. 2, respectively. First, the system modeling is done using the steps in Fig. 1. Then, using the steps in Fig. 2, the ball mill is controlled. The central parts of the proposed method include removing outliers with the adaptive quantum basis, extracting features from the audio signal using the Fourier transform, using the PCA method to select the central part, and modeling based on the least squares.

According to Fig. 1, the first step is acoustic signal acquisition. In the next step, removing outliers and noise is performed on the signal to reduce the modeling error. The proposed method to remove noise is as follows: first, a basis (dictionary) is created using the input signal. To separate the primary information and noise, the desired signal is projected onto the dictionary. After obtaining the coefficients in the dictionary domain, the denoised signal is reconstructed. In the next step, the Fourier transform of the denoised signal is calculated. Then, several essential features are extracted from the Fourier transform. To reduce the dimensions of the features, the PCA-based method is used in the next step. Finally, using the least squares method, the relationship between particle size distribution and acoustic signal is obtained as a system model. In Fig. 2, the ball mill system is controlled using the model obtained at this step. In the following sections, the different steps will be explained.

A. Denoising Scheme

Since ball mills are applied in industrial environments, the received acoustic signals contain noise and outlier points. One of the most critical parts of the ball mill control is the noise removal step. If the acoustic signal is not correctly denoised, it can cause errors in the control system. There are different methods to remove outliers and noise in the signal. Signal decomposition is one of the most common methods of noise removal. For this purpose, methods based on Fourier, Wavelet, and SVD are usually used [13-15]. These methods use a fixed basis to represent the signal and noise removal. In this paper, we used the denoising method, which is based on an adaptive basis. Recently, quantum computing has been used in various fields. Reconstructing geological images and removing noise is one of its most important applications [16-18]. In this article, the signal is denoised by changing in the idea proposed in [18].

In the proposed method, the discrete version of the Schrodinger equation in quantum computation is used to build an adaptive basis or dictionary. Therefore, this method has the advantage of calculating the transformation compatible with the desired acoustic signal. Another reason for using this method is that it can eliminate most common noises in industrial environments, because of its adjustable parameters [18]. It should be noted that in this study, we have used the idea of [18], only to create a dictionary**.** The rest of the algorithm, which concerns noise elimination, is the innovation of this article.

B. Construction of Quantum Adaptive Basis (Dictionary)

The stationary Schrodinger equation in quantum mechanics (a quantum particle with energy E and in potential $V(r)$) is as follows:

$$-\frac{\hbar^2}{2m}\nabla^2\varphi = -V(r) + E\varphi \quad (1)$$

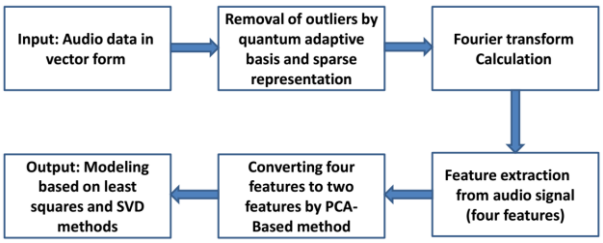


Fig. 1. Modeling flowchart of the proposed method

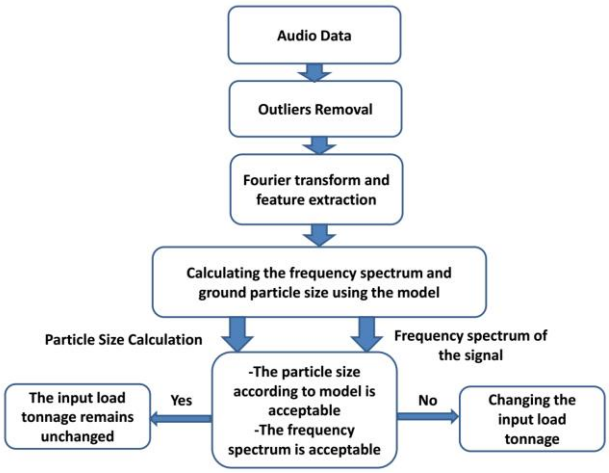


Fig. 2. Control flowchart of the proposed method

where m is the quantum particle mass, $\hbar/2\pi$ is the reduced Planck's constant, and ∇^2 is the Laplacian operator. This equation determines the wave function $\varphi(r)$ regarding spatial coordinates. To use this equation (in making the dictionary and signal denoising), the desired signal is considered a potential $V(r)$. Assuming a constant potential and a wave function with a periodicity of L, the answer to the Schrodinger equation will be in the following form:

$$\varphi(r) = Aexp(i\frac{\sqrt{2m(E-V)}}{\hbar}r) \tag{2}$$

where A is the amplitude. According to (2), by discretizing the space with N values, we will get N answers, and the frequency of the wave function will depend on $\sqrt{E-V}$. To use this equation to construct an adaptive basis, the desired signal is replaced by the potential $V(r)$. Therefore, it is necessary to create the discrete version of the Schrodinger equation. In operator form, (1) will be as follows:

$$H\varphi = E\varphi. \tag{3}$$

where $H = -\frac{\hbar^2}{2m}\nabla^2 + V$ is called the Hamiltonian operator. By replacing $y(t)$ (acoustic signal) with the size of N instead of $V(r)$, the size of H will be $N \times N$. When the discrete definitions of derivative and Laplacian are inserted, (3) will be as follows:

$$(\nabla\varphi)_t = \varphi(t+1) - \varphi(t)$$
$$(\nabla^2\varphi)_t = \varphi(t+1) - 2\varphi(t) + \varphi(t-1)$$

$$\text{Thus } (H\varphi)_t = -\frac{\hbar^2}{2m}(\varphi(t+1) - 2\varphi(t) + \varphi(t-1) + y(t)\varphi(t)) \Downarrow (H\varphi)_t = \left(y(t) + 2\frac{\hbar^2}{2m}\right)\varphi(t) - \frac{\hbar^2}{2m}\varphi(t+1) - \frac{\hbar^2}{2m}\varphi(t-1). \tag{4}$$

The simpler form of (4) is as follows:

$$(H\varphi)_t = \sum_{s=t-1}^{t+1} H(t,s)\varphi(s) \text{ for } t = 1,2,3,\dots,N. \Downarrow H(t,s) = \begin{cases} y(t) + 2\frac{\hbar^2}{2m} & for\ t = s \\ -\frac{\hbar^2}{2m} & for\ s = t+1\ and\ t = s-1 \\ 0 & otherwise \end{cases} \tag{5}$$

Therefore, the Hamiltonian matrix will be in the following form:

$$H = \begin{bmatrix} y(1) + 2\frac{\hbar^2}{2m} & -\frac{\hbar^2}{2m} & 0 \\ -\frac{\hbar^2}{2m} & \ddots & -\frac{\hbar^2}{2m} \\ 0 & -\frac{\hbar^2}{2m} & y(N) + 2\frac{\hbar^2}{2m} \end{bmatrix}. \tag{6}$$

According to (3), eigenvectors of the Hamiltonian matrix can be computed for different values of E, and they form a basis in the Hilbert space. This matrix has N eigenvectors, which are represented by φ_i. These eigenvectors form a basis, which is called a dictionary in this article. It should be noted that one of the problems of the Schrodinger equation is the localization of the basis, which is caused by the Anderson localization phenomenon [19]. To solve this problem, the input signal is first smoothed using a Gaussian filter [18]. Then, the dictionary is made from the smoothed signal. The desired audio signal will have a sparse representation in this dictionary. Using this method, the noisy and non-noisy parts are separated, and the signal will be denoised. In the following, the idea of sparse representation will be expressed.

C. Sparse Representation Model

Recently, the sparse model has been used in various applications such as reconstruction and noise removal [20-21]. In this model, $\boldsymbol{\varphi} = [\varphi_1, \dots, \varphi_N]$ represents the dictionary, φ_i is the dictionary atom with the unit norm, and $\boldsymbol{y}$ represents the vectorized signal. In sparse representation, it is assumed that the desired signal can be represented by a small number of dictionary atoms. The $\boldsymbol{y}$ signal can be represented as follows:

$$\boldsymbol{y} = \boldsymbol{\varphi}\beta = \sum_{i=1}^{N} \varphi_i \beta_i \tag{7}$$

where β will be the sparse representation of $\boldsymbol{y}$ in $\boldsymbol{\varphi}$, if it contains a few non-zero coefficients. The problem of finding β is formulated as follows:

$$\beta = argmin_{\beta} \quad \|\boldsymbol{y} - \boldsymbol{\varphi}\beta\|_2^2 \quad \text{subject to} \quad \|\beta\|_0 \leq s \qquad (8)$$

where $\|\beta\|_0$ means the number of non-zero elements of β and s is sparsity. Matching Pursuit (MP) method is commonly used to solve the above problem [20]. This method starts with finding the atom that best matches the signal. The error in this step is defined as follows:

$$E_i = min_{\alpha}\|\alpha\varphi_i - \boldsymbol{y}\| \qquad (9)$$

where α is a scalar. The above problem is solved using $\alpha^* = \langle \boldsymbol{y}, \varphi_i \rangle$, and the index of the best match is calculated as i_0. Then, the residual signal $\boldsymbol{r}$ is calculated, and this process continues until the desired accuracy is reached. The MP algorithm to achieve a sparse representation of a signal is shown in Table 2.

After calculating the sparse representation of the signal in the dictionary space, the denoised signal is obtained using $\hat{\boldsymbol{y}} = \boldsymbol{\varphi}\beta$. The steps of the proposed denoising algorithm are shown in Table 3. According to (6), the size of the Hamiltonian matrix is $N \times N$. If the size of the input signal is large, the computational complexity in dictionary construction will be very high. Therefore, in these cases, the input signal is divided into different patches, and the algorithm of Table 3 is executed on each patch. In addition to reducing the computational complexity, this method will also solve the problem of lack of memory. It should be noted that in this article, the parameters are set as $\frac{\hbar^2}{2m} = 0.4,\ \eta = 0.2$, and $\sigma = 5$.

TABLE 2
MATCHING PURSUIT ALGORITHM

1. $\boldsymbol{r} = \boldsymbol{y}, \beta^1 = 0,\ z = 1$
2. ***while*** $(\boldsymbol{r} > \eta)$ ***do***
3. $E_i = min_{\alpha}\|\alpha\varphi_i - \boldsymbol{y}\|$
4. $i_0 = argmin_{\ i} E_i$
5. $\beta_i^z = \begin{cases} \beta_i^{z-1}: & i \neq i_0 \\ \beta_i^{z-1} + \langle \boldsymbol{r}, \varphi_i \rangle: & i = i_0 \end{cases}$
6. $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{\varphi}\beta^z$
7. $z = z + 1$
8. ***end while***
9. $\beta = \beta^{z-1}$

TABLE 3
PROPOSED DENOISING ALGORITHM

Input: $\boldsymbol{y}$ *(Noisy Signal),* $\frac{\hbar^2}{2m}$, η, σ

1- *Smooth the input acoustic signal with a Gaussian filter*
2- *Compute the Hamiltonian matrix using (6)*
3- *Calculate the eigenvectors of the* H *matrix (get* φ_i*)*
4- *Consider the* φ_i *vectors as a dictionary of atoms and normalize them (get* $\boldsymbol{\varphi}$*)*
5- *Get sparse representation* β *using the MP algorithm (Table 2)*
6- *Reconstruct the denoised signal using* $\hat{\boldsymbol{y}} = \boldsymbol{\varphi}\beta$

Output: $\hat{\boldsymbol{y}}$ *(Denoised Signal)*

D. Fast Fourier Transform and Feature Extraction

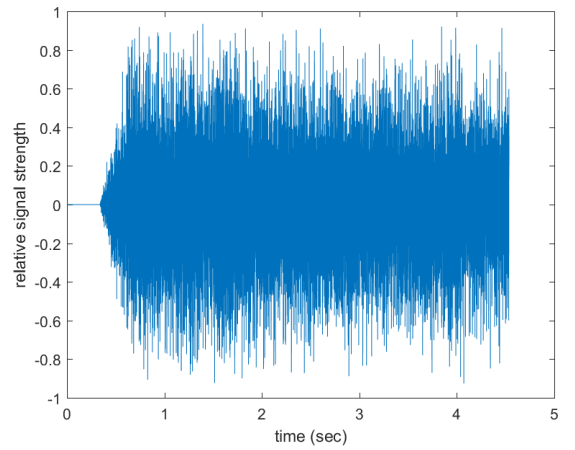
Fast Fourier transform is a method that computes the DFT (Discrete Fourier Transform) of a sequence [22]. It reduces the computations needed for N points from $2N^2$ to $2Nlog_2N$. Fourier analysis converts a signal from its original (often time) to the frequency domain. In this part, the fast Fourier transform of the denoised audio signal and the power spectrum are calculated. In Fig. 3, one sample of acoustic signal at different sampling times is shown along with their frequency spectrum. According to Fig. 3, at each sampling time, one component of the power spectrum is considerably different energy than the others. Therefore, the amplitude and frequency of this component are selected as the feature extracted from this sample. Two essential features of the time domain, i.e., maximum and minimum sound intensity, are also used as extracted features. Therefore, four features extracted from the acoustic signal are considered candidates. In the next step, by reducing the dimensions of the features, they are converted into two main features.

E. Dimension Reduction Method

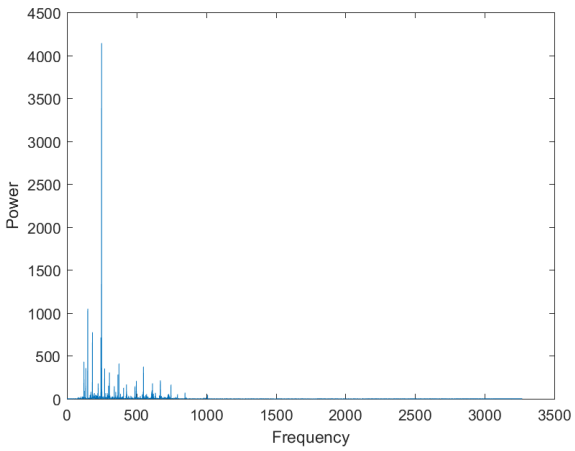
One essential tool to reduce the data dimensions is representation in principal component space [23]. Since the goal of this article is to model the ground particle size versus sound characteristics, data dimensions should be reduced after extracting the features. During the operation of the ball mill, the output product is sampled at different time intervals, and the audio signal is also recorded simultaneously. After sieve analysis, a particle size is assigned to each sampling time. Considering that the values of the variables were extracted from the sound at different sampling times, the data matrix size will be $M \times N$, where M is the number of observations, and N is the number of extracted features. To reduce the dimensions of the extracted features according to Table 4, the covariance matrix of the data is first calculated. Then, the eigenvectors of the covariance matrix are extracted. This matrix is called V. Finally, new data is formed by the product of X and V. In the next step, these data are used for modeling. Using this method, the principal components of the data are computed, and the dimension of the features is reduced.

F. Modeling Based on Least Squares Regression

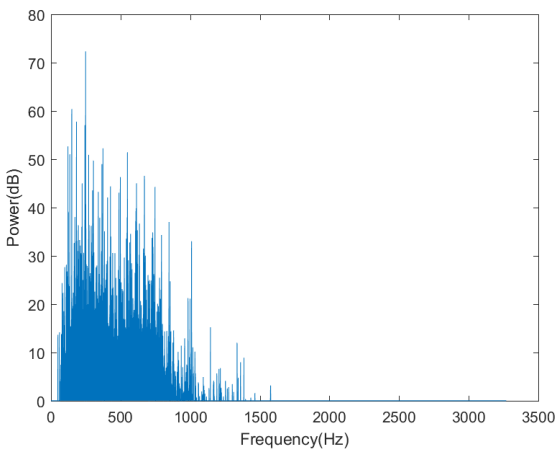
In this part, the method of least squares in identifying the ball mill system is briefly reviewed. This method, while being simple, is good enough to control the ground particle size and ball mill load based on practical tests.



(a)



(b)



(c)

Fig. 3. An example of an acoustic signal with a length of 4.5 seconds. (a) signal in the time domain, (b) Power spectrum components, (c) Power spectrum in decibels (dB)

The sum of the errors caused by the measurement and definition of the structure is indicated by e_t. In the least squares method, the goal is to estimate the parameters so that the sum of squared error ($\sum {e_t}^2$) is minimized [24].

TABLE 4
DIMENSION REDUCTION ALGORITHM

Input: Data Matrix $X_{M\times N}$

1- *Compute the covariance of the matrix* $X_{M\times N}$ *:* $covarianceMatrix = cov(X)$
2- *Extract the eigenvectors of the covariance matrix:* $[V, D] = eig(covariance\ Matrix)$
3- *Show the new data in principal component space:* $Y = X * V$

Output: Data Matrix $Y_{M\times p}$ *in Principal component space where* $p < N$

In this method, a linear structure for the relationship between input and output is in the following form:

$$y_t = u_t^T\theta + e_t \tag{10}$$

where y_t is the output at t, θ is vector parameters, and u_t is a vector containing system information. In this method, $S = \sum {e_t}^2$ must be minimized to calculate θ. Therefore, the linear regression equation for all sampling times is as follows:

$$\begin{cases} y_1 = u_1^T\theta + e_1 \\ y_2 = u_2^T\theta + e_2 \\ \vdots \\ y_N = u_N^T\theta + e_N \end{cases} \tag{11}$$

The equation (11) will be in the following form:

$$y = U.\theta + e;\ \begin{cases} y_{N\times 1} = [y_1\ y_2\ \dots\ y_N]^T \\ U_{N\times p} = [u_1\ u_2\ \dots\ u_N]^T \\ e_{N\times 1} = [e_1\ e_2\ \dots\ e_N]^T \\ \theta_{p\times 1} = [\theta_1\ \theta_2\ \dots\ \theta_p]^T \end{cases} \tag{12}$$

Therefore, $\hat{\theta}$ (estimated parameters) is calculated as follows:

$$S = \sum_{t=1}^{N} {e_t}^2 = [y - U.\theta]^T[y - U.\theta] = y^Ty - \theta^TU^Ty - y^TU\theta + \theta^TU^TU\theta$$
$$\Downarrow$$
$$\frac{\partial S}{\partial \theta} = -2U^Ty + 2U^TU\theta \tag{13}$$
$$\Downarrow$$
$$\hat{\theta} = (U^TU)^{-1}.U^Ty.$$

Inverting (U^TU), when $\det(U^TU)$ is close to zero, can cause errors in $\hat{\theta}$ calculation [25]. In this case, it is called an ill-posed matrix. Therefore, alternative methods should be used to calculate the inverse of (U^TU). In this article, the singular value decomposition (SVD) method is used to solve this problem. The summary of the SVD method is shown in Table 5. The singular value decomposition of U is as follows [15]:

$$U_{N\times p} = P_{N\times N}R_{N\times p}Q_{p\times p} \tag{14}$$

where P and Q are orthonormal matrices with the following characteristics:

$$P^TP = PP^T = I_{N\times N}$$
$$Q^TQ = QQ^T = I_{p\times p} \tag{15}$$

And matrix R is in the following form:

$$R = \begin{bmatrix} \sigma_{p\times p} \\ 0_{(N-p)\times p} \end{bmatrix}_{N\times p} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p \\ & & \mathbf{0} & \end{bmatrix}_{N\times p} . \quad (16)$$

Equation 12 can be written in the following form:

$$P^T.(y = U.\theta + e) \Rightarrow P^T y = P^T U\theta + P^T e$$
$$U = PRQ \Rightarrow P^T y = P^T PRQ\theta + P^T e \quad (17)$$

Where $PP^T = I$ and assuming $y^* = P^T y$, $e^* = P^T e$, $\theta^* = Q\theta$, (17) can be written in the following form:

$$y^* = R\theta^* + e^*. \quad (18)$$

To calculate θ^*, equation (18) is expanded as follows:

$$\begin{bmatrix} y_1^* \begin{Bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_p^* \end{Bmatrix} \\ y_2^* \begin{Bmatrix} y_{p+1}^* \\ \vdots \\ y_N^{*N} \end{Bmatrix} \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_1^* \\ \theta_2^* \\ \vdots \\ \theta_p^* \end{bmatrix} + \begin{bmatrix} \begin{Bmatrix} e_1^* \\ e_2^* \\ \vdots \\ e_p^* \end{Bmatrix} e_1^* \\ \begin{Bmatrix} e_{p+1}^* \\ \vdots \\ e_N^* \end{Bmatrix} e_2^* \end{bmatrix}. \quad (19)$$

$$\Downarrow$$

$$\begin{cases} y_1^* = \sigma_{p\times p}\theta^* + e_1^* \\ y_2^* = e_2^* \end{cases}$$

Therefore, $\theta^* = (\sigma^{-1} y_1^*)$ and using $\hat{\theta} = (Q^T\theta^*)$, $\hat{\theta}$ is calculated. Consequently, for modeling using the least squares method, the information matrix U is first calculated at different sampling times. The model parameters are estimated using the SVD method.

III. Practical Tests and Results

In this part, the experiments for modeling purposes were designed based on the factorial method and relative response time [26]. According to the previous description, for modeling the relationship between ground particle size and acoustic signal, before and after ball charging, the output (ground particles) of the ball mill is sampled at different times, and the acoustic signal is also recorded. Each sample is dried in the laboratory, and the sieve analysis is performed. In this section, the ASTM (American Standard Test Sieve Series) series and the Wet Sieve Analysis method are used for PSD (Particle Size Distribution) analysis. D_{80} is a characteristic diameter extracted from PSD and will be used in this section. Therefore, in the laboratory, 200 grams of the sample was separated, and the rest was stored as an archive. The representative sample was sieved on several different sieves. The material left on the first sieve should not be less than 10% of the total sample. If there is no load on it, a smaller sieve should be selected.

TABLE 5
SVD METHOD

Input: Matrix U

1- *Calculate the SVD of* U: $U = PRQ$.
2- *Calculate* $y^* = P^T y$ *and get* y_1^*.
3- *Calculate* $\theta^* = (\sigma^{-1} y_1^*)$.
4- *Calculate* $\hat{\theta} = (Q^T\theta^*)$.

Output: $\hat{\theta}$

Then, the amount of the remaining materials on each sieve is weighed, and the numbers are placed in a table along with the sieve number. Finally, graphs are drawn based on the data of the 80% of the material passed through the sieve, and the results are interpreted. Fig. 4 shows the sampling and analysis steps, respectively, in the plant and the laboratory. In the next part, the optimal load of the ball mill will be selected through sieve analysis for a better release of minerals.

A. Sieve Analysis Results

For a better release of each mineral, it is necessary to grind the input load to a specific size. In the studied plant, this value is in the range of 110 to 120 microns. Therefore, the feed rate is optimal when the ball mill can grind it to the intended size range. This range is specific to this factory, due to the specifications of its available facilities, and can be different in other places. The ground particle size results for different feed rates, before and after ball charging, are shown in Table 6, Table 7, and Fig. 5. Feed rate changes have a direct effect on retention time. If the feed rate increases, the retention time will decrease. With the rise in the feed rate, the retention time will be reduced, and less grinding will be done on the materials. Consequently, the output particle size will be coarser. According to the tests done in this part, the retention time was estimated to be between 8 and 15 minutes. Therefore, the sampling time of 42 minutes was selected based on the Nyquist theorem and retention time [27]. According to the Nyquist theorem, the signal sampling frequency should be chosen so that the vital signal data cannot be disturbed. During this time, it is expected that the output will reach stability in all different input tonnages. Also, all the data resulting from the grinding steps and fluctuations (used in modeling) can be seen in the output samples. According to Fig. 5, before the ball charging, a peak is observed in the 20t/h graph. It shows that this cargo is relatively large for the ball mill available in the factory, and the power of the ball mill to grind the load is small.

As a result, the output of the ball mill in the mentioned tonnage has a coarser particle size than other tonnages; therefore, the value of 20t/h is the critical value. Also, it can be concluded that the ball mill shows a better grinding performance at low tonnages.

Fig. 4. Sampling and sieve analysis steps

Therefore, the corresponding ground particle size curves are always downward over time. This process continues until the particles soften, and gradually, the curves become horizontal. Unlike the other charts, the 10t/h chart has less fluctuation and smaller particle size. As a result, it is better for the release of minerals. After charging new balls, the changes in all input tonnages decreased. In the previous chart, the feed rate of 20t/h is considered a critical feed rate due to the coarseness of the output particles even after the stability time. Compared to the other charts, the 10t/h chart has less fluctuation and a smaller particle size. Therefore, according to the analysis and requested size range, the optimal feed rate for the ball mill is 10t/h.

B. Spectral Analysis Results

In this part, to investigate the sound signal behavior at different feed rates, the signal power spectrum was calculated using (20) and (21). In the mentioned equations, $Y(k)$ is the FFT of $Y(t)$ (acoustic signal), and N is the signal dimension. Also, equation (21) calculates the power in decibels (dB).

$$Power = \frac{1}{N}\sum_{k=1}^{N}|Y(k)|^2 \tag{20}$$

$$Power(dB) = 20log_{10}(\frac{1}{N}\sum_{k=1}^{N}|Y(k)|^2) \tag{21}$$

Fig. 6 shows the power spectrum for three different feed rates, including 5t/h, 10t/h, and 20t/h. In this figure, the top row shows the components of the power spectrum, and the bottom row is the power spectrum in decibels (dB). According to Fig. 6, in the feed rate of 5t/h, the power spectrum is in a higher frequency range than the other two feed rates. In this load, the amplitude of power components is also more extensive than the other two loads. The reason is that when the feed rate is low, balls hit the mill body more frequently, and consequently, the intensity and frequency of the produced sound will be higher.

TABLE 6
PSD ANALYSIS (D_{80} RESPONSE BEFORE BALL CHARGING (MICRON))

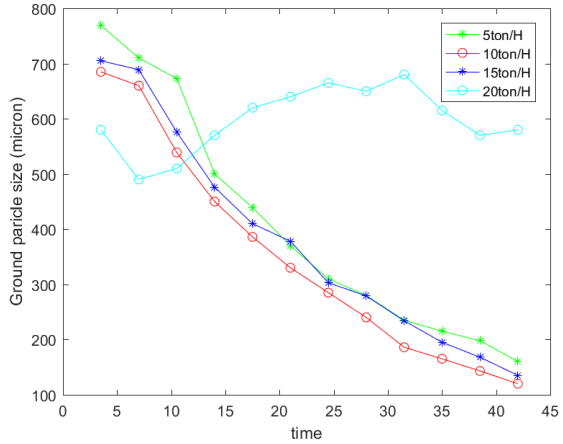
	Feed Rate			
Time (min)	5t/h	10t/h	15t/h	20t/h
3.5	769	685	705	580
7	710	660	689	490
10.5	673	539	576	510
14	500	450	476	570
17.5	439	386	410	620
21	370	330	378	640
24.5	310	285	303	665
28	280	240	279	650
31.5	235	186	234	680
35	215	165	195	615
38.5	198	143	168	570
42	160	120	135	580

TABLE 7
PSD ANALYSIS (D_{80} RESPONSE AFTER BALL CHARGING (MICRON))

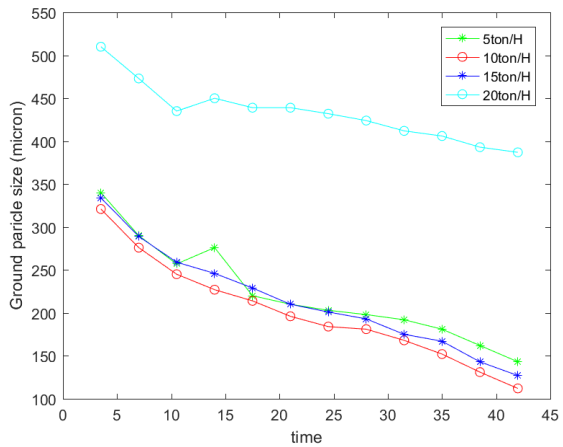
	Feed Rate			
Time (min)	5t/h	10t/h	15t/h	20t/h
3.5	340	321	334	510
7	290	276	289	473
10.5	257	245	259	435
14	276	227	246	450
17.5	220	214	229	439
21	210	196	210	439
24.5	203	184	201	432
28	198	181	193	424
31.5	192	168	175	412
35	181	152	167	406
38.5	162	131	143	393
42	143	112	127	387

In the feed rate of 20t/h, due to the high load of the machine, the sound resulting from the collision of the balls with the load is muffled. Therefore, in this case, the frequency of the power spectrum is shallow, and the amplitude is less than in the previous point. In the feed rate of 10t/h, the signal's frequency is in the middle band, and the amplitude of the power spectrum is also lower. This means that the mill is working normally, and the frequency and intensity of the sound are also suitable. Therefore, the summary of this part of the studied ball mill is as follows:

- The frequency range of the power spectrum for all cases is from 100 to 1300 Hz.
- If the feed rate is low, the sound power spectrum will be at higher frequencies, which is in the range of 700 to 1300 Hz.



(a)



(b)

Fig. 5. PSD Analysis (Ground particle size (D_{80}) graph).
(a) before ball charging, (b) after ball charging

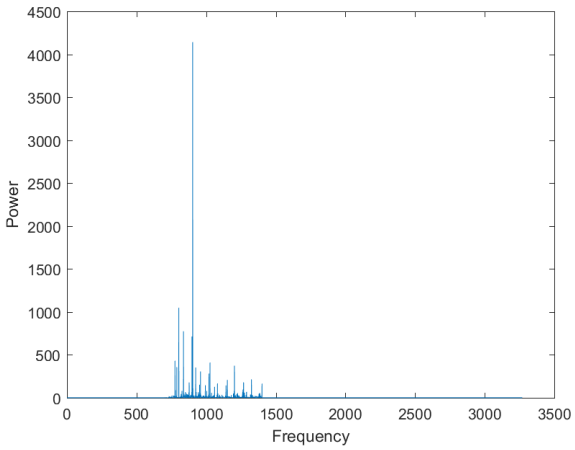
- If the feed rate is high (critical), the sound power spectrum will be at low frequencies, which is in the range of 100 to 300 Hz.
- If the feed rate is standard, the sound power spectrum will be in the middle frequencies and the range of 300 to 700 Hz.

Therefore, according to the results and the frequency range obtained for the sound, the feed rate of the ball mill can be divided into three states: standard feed rate, low feed rate, and critical feed rate.

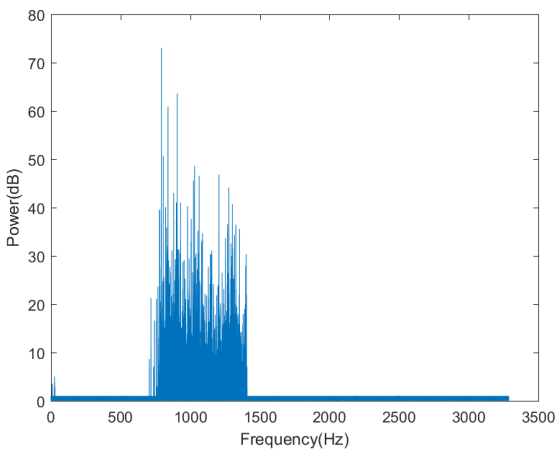
C. Modeling Based on Least Squares

As mentioned in the previous sections, four features (including the maximum and minimum sound intensity in the time domain, and the amplitude and frequency of the maximum power spectrum components) are extracted as candidates from the acoustic signal in different time intervals.

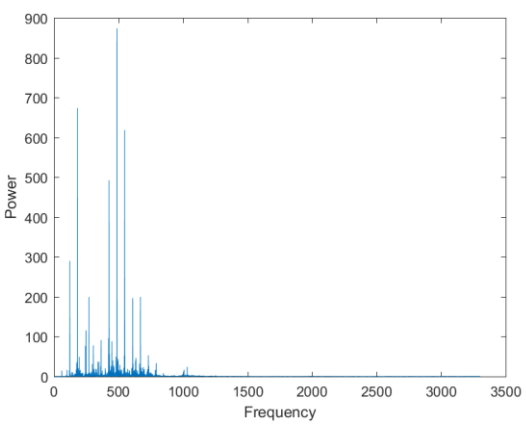
According to Table 4, two features (i.e., the maximum and minimum sound intensity in the time domain) were ignored due to their low change interval. The reason is that to reduce the dimensions of the features (Table 4), the PCA-based method has been used. In this method, characteristics with more changes are considered more important in modeling. On the other hand, according to the obtained results, the range of changes in maximum and minimum sound intensity is tiny. The remaining two features used for modeling included the amplitude (A_m) and frequency (f_m) of the most significant component of the power spectrum.



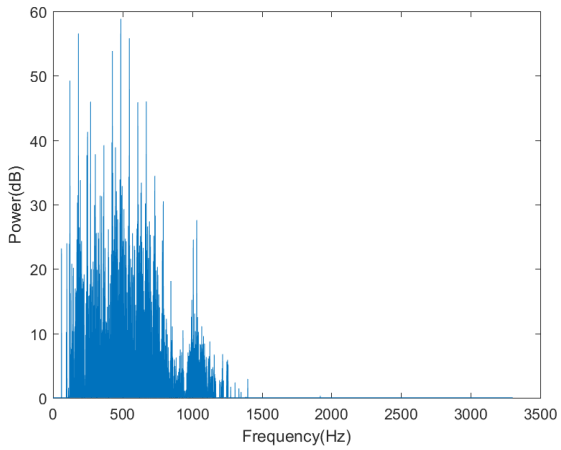
(a)



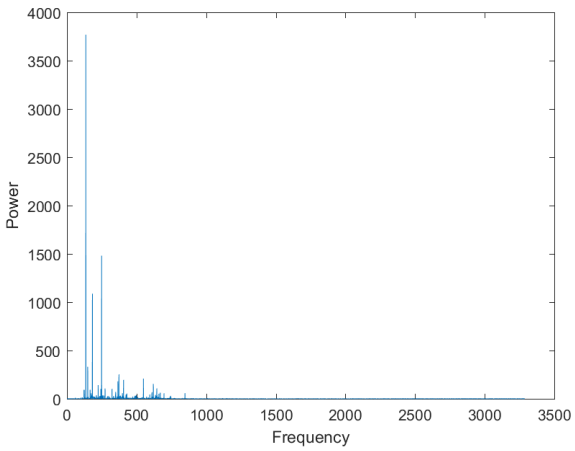
(b)



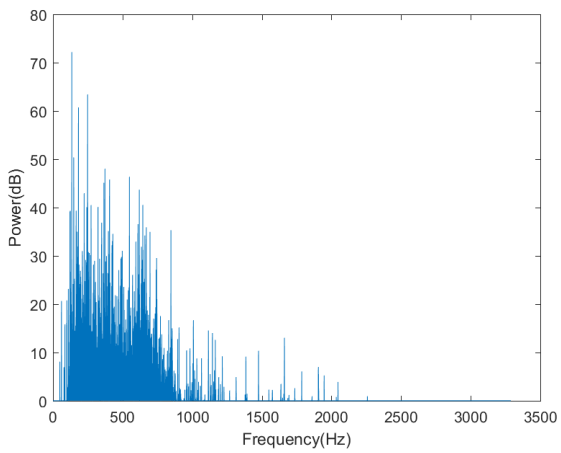
(c)



(d)



(e)



(f)

Fig. 6. Acoustic power spectrum at different feed rates (a),(b) low feed rate (5t/h) , (c),(d) standard feed rate (10t/h), (e),(f) high feed rate (20t/h)

In this article, the goal is to find a model that shows the relationship between the ground particle size and the audio signal. To achieve this goal, A_m and f_m are extracted from sound in all sampling time intervals. Then, the ratio f_m/A_m is innovatively used as a feature extracted from the sound. Sound is the result of vibration. This ratio can physically express the amount of vibration resulting from the collision of the balls with the ore. Using this feature, critical and standard conditions of the mill can be identified. According to the amount of vibration created in the ore, the normal and vital conditions of the mill will be determined, and a specific sound frequency range will be produced. In section 3.1, the ground particle size for all samples has been extracted using sieve analysis. Therefore, according to section 2.4, a mathematical model can be obtained by putting the ground particle size on the Y-axis and the feature value on the X-axis. In this study, the least squares method based on polynomials has been used.

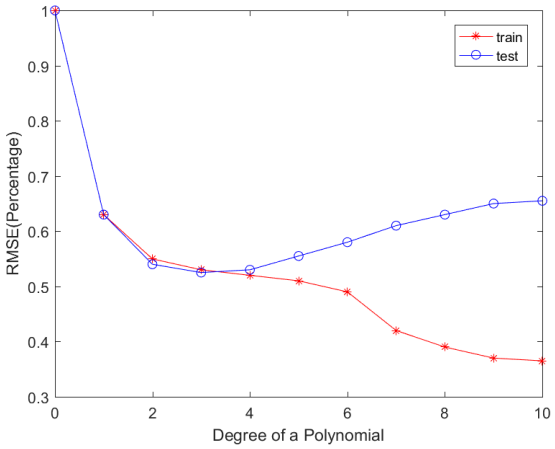
The polynomial degree selection is one of the primary challenges in modeling. It should be noted that the purpose of modeling is not to predict the output for the given data; instead, it is to predict the outcome of the unseen data. A model that makes reasonable predictions on new unseen data has generalization ability. A simple and effective method to determine if a model can be generalized is to divide the original data into a training set and a test set. Therefore, 80% of the data is used for training and 20% for model testing. After polynomial modeling with different degrees, the value of RMSE is calculated using (22). In this equation, Y_i is the actual value of the data, and $\hat{Y}_i$ is estimated by the model.

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2 \tag{22}$$

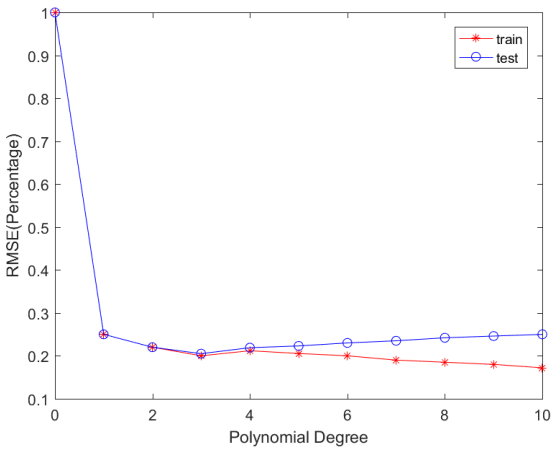
Then, the RMSE diagram is drawn for the training and test data in different degrees in polynomial equations. Finally, the degree with the lowest RMSE value in the test data is selected as the most preferable one. In Fig. 7, the model validation is done for the data before and after the ball charging. According to this figure, degrees 2, and 3 are reasonable choices. Given the trade-off between simplicity and accuracy, degree 2 is chosen for modeling.

The modeling results for the feed rate of 10t/h before and after ball charging are shown in Fig. 8 and Table 8. Using this model, the ground particle size is estimated based on the feature value extracted from the sound. Therefore, the control plan (Fig. 2) of the ball mill will be as follows:

- First, the optimal feed rate (i.e., 10t/h) is set, based on the previous results.
- Then, the mill sound signal is recorded every 3.5 minutes, after the retention time (15min).
- After the noise removal process, the features and the sound power spectrum are extracted.



(a)



(b)

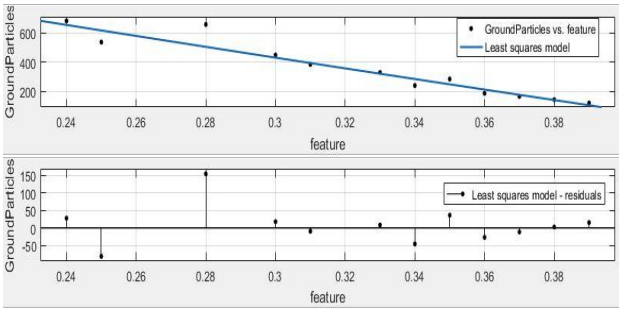
Fig. 7. Model validation chart
(a) before ball charging, (b) after ball charging

- Using the extracted features, the ground particle size is calculated.
- If the calculated ground particle size is 10% different from the desired one and the sound power spectrum is not in the middle-frequency range, the feed rate value should be changed.

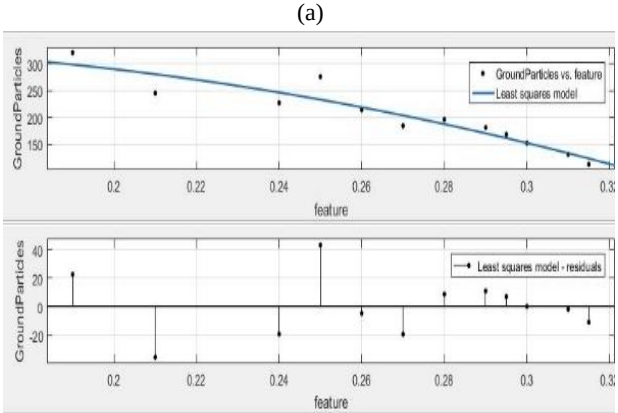
D. Built-in Control Device

Fig. 9 shows installing the microphone and adjusting the feed rate to the ball mill in the Lakan lead-zinc processing plant. The microphones are installed in a place closest to the bottom of the mill. In this article, a BY-M1 BOYA condenser microphone was used for recording the sound.

Due to the high quality and sensitivity of the condenser microphones, more details of the signal are recorded, and the sampled signal can be used to diagnose the working conditions of the ball mill.



(a)



(b)

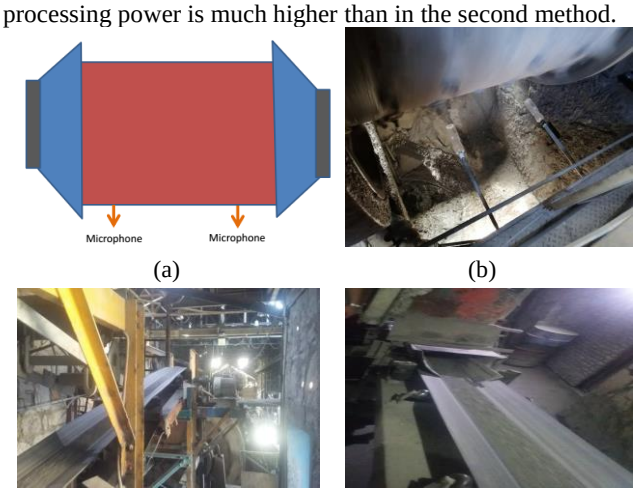
Fig. 8. Modeling based on least squares before and after ball charging in feed rate 10t/h
(a) before ball setting, (b) after ball charging

TABLE 8
THE IDENTIFIED MODEL IN FIG. 8

Before Ball Charging: $f(x) = p_1x^2 + p_2x + p_3, \quad p_1 = 1.859,\ p_2 = -182.1,\ p_3 = 339.6$ RMSE= 54.97
After Ball Charging: $f(x) = p_1x^2 + p_2x + p_3, \quad p_1 = -4817,\ p_2 = 1029,\ p_3 = 276.9$ RMSE= 22.99

The control plan proposed in this article (Fig. 2) can be implemented in two ways. One is to use a computer and connect a microphone to it so that the received signal can be analyzed using the proposed method in MATLAB software. Using this method, the ground particle size is calculated, and the feed rate is changed if needed. In the other way, a unique device is made using an Arduino microcontroller (Fig. 10 shows different parts of the manufactured device). In the mentioned circuit, the received audio signal is first amplified using an analog amplifier. In the next step, the signal is transferred to the microprocessor. Then, the signal is denoised, and features are extracted. After that, the ground particle size is estimated using the extracted models in Fig. 8, and the sound power spectrum is also calculated. If the ground particle size and sound power spectrum condition are not ideal (according to 3.2), the buzzer starts making sounds, and the control switch

is turned on to change the feed rate. In this article, both methods were practically implemented. In the first method, the cost of implementation and construction is high, but the data processing power is much higher than in the second method.



(a) (b)

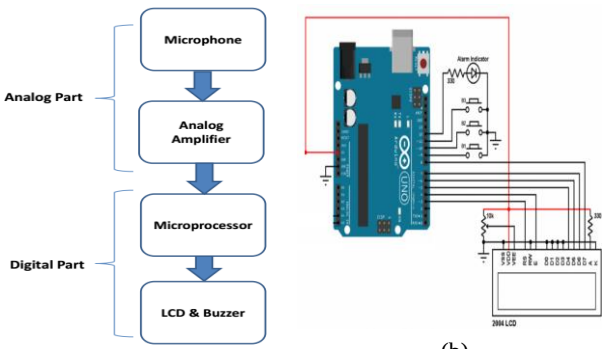
(c) (d)

Fig. 9. Installing the microphones near the mill floor and adjusting the input feed rate. (a) Location of microphones, (b) Microphone installation in the plant, and (c), (d) Setting the feed rate

However, in the second method, the manufactured device is portable and, hence, cost-effective due to its small dimensions. The fabricated device can reduce the energy consumption of the ball mill. Also, by adjusting the optimal feed rate, the average size of ground products can be decreased and set in the appropriate range. In the proposed control system, since the ball mill load is standard, ball mill energy consumption will be optimal. Optimal load control will also increase the ball charging time. Therefore, a smaller number of balls will be used. Consequently, the system designed in this article will increase the efficiency of the ball mill in practice. As explained in the introduction section, the studied ball mill liner is made of plastic. Metal liners produce a lot of noise, and distort the sound resulting from the collision of the balls with the particles. But plastic liners do not make noise due to their elasticity. Therefore, the received sound will result from the impact of the material with the balls. Hence, one of the limitations of the proposed method is that the liners must be plastic to minimize the error. Because of the fact that nowadays, plastic liners are primarily used in ball mills, the presented method will not have severe limitations. One of the most essential points in the designed system is repeatability. To achieve it, the identified model should be updated as the system parameters change. For this purpose, in the new device version, the modeling process is done in different time frames. Using this method, the model is continuously updated. To do so, the output ground particles are analyzed using a CAMSIZERX2 device. Particle size distribution data (i.e., D_{80}) is sent to the device online. Then, the modeling process is done based on the least square error method, and the model will be updated.

IV. Conclusion

In this study, an innovative combined method was implemented to control the ground particle size and load of the ball mill. In this method, the relevant modeling was first done using both the acoustic signal and the ground particle size diagram.



(a) (b)

(c)

Fig. 10. Prototype of the ball mill control device. (a) schematic of the device, (b) a digital circuit, (c) a prototype of the manufactured device

To do so, the output (ground particles of the ball mill) was first sampled, and the acoustic signal acquisition process was also performed. After the sieve analysis, the ground particle size diagram was drawn. In the next step, using the new method of quantum adaptive basis, the acoustic signal outliers were removed. Then, the Fourier transform of the signal and the power spectrum were calculated. Four features, namely the maximum and minimum values of the sound in the time domain, and the maximum of the power spectrum component (amplitude and frequency), were extracted from the acoustic signal. In the next step, a PCA-based method was applied to reduce the dimension of the features. Finally, the regression model between sound signal and ground particle size was obtained using the least squares method. This model is also used to control the ball mill. The resulting model can detect the ground particle size and load status (standard, low, high) by recording the sound of the ball mill, when in operation. In this article, all the steps of the proposed method were practically tested on the ball mill of the Lakan lead-zinc processing plant. Also, a device that controls the ground particle size and ball mill load was built. According to the results, the optimal feed

rate for the ball mill is 10t/h. In this case, the size of the ground particles is suitable (110-120 microns), and the sound power spectrum is 300-700 Hz. It should be noted that the optimal feed rate obtained in this article is only specific to the studied ball mill. That is, the optimal feed rate in other factories might be different, and it must be calculated using the proposed scheme. According to the analysis and results of the study, the proposed control plan will reduce the energy consumption of the ball mill and increase the ball charging time. Therefore, the proposed method will increase efficiency in this factory. The authors of this article aim to measure the amount of energy and the number of balls saved in this process for their future research**.** Also, the modeling of more complex ball mill systems will be investigated in the following study.

References

[1] S. Singh, S. K. Singh, P. K. Harijan, S. K. Yadav, and A. Kumar, "Investigation on the effect of Fe impurity pickups during ball milling and Ni dispersion on the microwave absorption performance of ball milled Fe impurities-Ni/SiC composites," Journal of Materials Science: Materials in Electronics, Jul. 2022.

[2] A. Hajati, S. Shafaei, & M. Noaparast, "A novel approach to optimize grinding circuits-modelling strategy to monitor ball mill particle size distribution data at Lakan plant, " IJST, Transactions of Mechanical Engineering, vol. 35, pp.85-100, 2011.

[3] A. M. Tsirlin and A. I. Balunov, "Optimal Control of Technological Processes," Processes, vol. 11, no. 6, pp. 1835–1835, Jun. 2023.

[4] Y. Li, "Modelling tumbling ball milling based on DEM simulation and machine learning," (Doctoral dissertation, UNSW Sydney), 2023.

[5] K. S. Cho, S. Kim, and Young Hun Lee, "Correlation between Acoustic Intensity and Ground Particle Size in Alumina Ball Mill Process," Journal of The Korean Ceramic Society, vol. 55, no. 3, pp. 275–284, May 2018.

[6] J. Lv, Z. Wang, and S. Ma, "Calculation method and its application for energy consumption of ball mills in ceramic industry based on power feature deployment," Advances in Applied Ceramics, vol. 119, no. 4, pp. 183–194, Feb. 2020.

[7] Halit Sübütay and İlyas Şavklıyıldız, "Effect of High-Energy Ball Milling in Ternary Material System of (Mg-Sn-Na)," Crystals, vol. 13, no. 8, pp. 1230–1230, Aug. 2023.

[8] Shashi Prakash Dwivedi et al., "Effect of ball-milling process parameters on mechanical properties of Al/Al2O3/collagen powder composite using statistical approach," vol. 15, pp. 2918–2932, Nov. 2021.

[9] J. Tang, J. Qiao, Z. Liu, X. Zhou, G. Yu, and J. Zhao, "Mechanism characteristic analysis and soft measuring method review for ball mill load based on mechanical vibration and acoustic signals in the grinding process," vol. 128, pp. 294–311, Nov. 2018.

[10] J. Shi, G. Si, and Y. Zhang,"Application of Fractional Fourier Transform for Prediction of Ball Mill Loads Using Acoustic Signals," IEEE Access, Jan. 2019.

[11] X. Wang, K. Sun, H. Zhang, X. Wan, and C. Yang, "Mill Load Identification Method for Ball milling Process Based on Grinding Signal," IFAC-PapersOnLine, vol. 54, no. 21, pp. 7–12, Jan. 2021.

[12] Y. Li, J. Bao, T. Chen, A. Yu, and R. Yang, "Prediction of ball milling performance by a convolutional neural network model and transfer learning," Powder Technology, vol. 403, p. 117409, May 2022.

[13] S. Kalantari, M. Ramezani, A. Madadi, and V. V. Estrela, "Reduction AWGN from Digital Images Using a New Local Optimal Low-Rank Approximation Method," Smart innovation, systems and technologies, Dec. 2020.

[14] C. Tian, M. Zheng, W. Zuo, B. Zhang, Y. Zhang, and D. Zhang, "Multi-stage image denoising with the wavelet transform," arXiv (Cornell University), Sep. 2022.

[15] S. Kalantari, M. Ramezani, and A. Madadi, "Introducing a New Hybrid Adaptive Local Optimal Low Rank Approximation Method for Denoising Images," International Journal of Industrial Electronics Control and Optimization, vol. 3, no. 2, pp. 173–185, May 2020.

[16] S. Kalantari, A. Madadi, & M. Ramezani, "Reconstruction of Geological Images Based on an Adaptive Spatial Domain Filter: An Example to Introduce Quantum Computation to Geosciences," International Journal of Mining and Geo-Engineering, vol. 57, no. 2, pp. 183-194, 2023.

[17] Muhammad Sahimi and Pejman Tahmasebi, "The Potential of Quantum Computing for Geoscience," Transport in Porous Media, vol. 145, no. 2, pp. 367–387, Sep. 2022.

[18] S. Dutta, A. Basarab, Bertrand Georgeot, and D. Kouamé, "Quantum Mechanics-Based Signal and Image Representation: Application to Denoising," IEEE open journal of signal processing, vol. 2, pp. 190–206, Jan. 2021.

[19] P. W. Anderson, "Absence of Diffusion in Certain Random Lattices," Physical Review, vol. 109, no. 5, pp. 1492–1505, Mar. 1958.

[20] S. Kalantari and M. J. Abdollahifard, "Optimization-based multiple-point geostatistics: A sparse way," Computers & Geosciences, vol. 95, pp. 85–98, Oct. 2016.

[21] H. Liu, J. Zhang, and R. Xiong, "CAS: Correlation Adaptive Sparse Modeling for Image Denoising," IEEE transactions on computational imaging, vol. 7, pp. 638–647, Jan. 2021.

[22] E. Rajaby and S. M. Sayedi, "A structured review of sparse fast Fourier transform algorithms," Digital Signal Processing, vol. 123, p. 103403, Apr. 2022.
[23] B. Hasan, B. M. S., & A. Abdulazeez, "A review of principal component analysis algorithm for dimensionality reduction," Journal of Soft Computing and Data Mining, vol. 2, pp. 20-30, 2021.
[24] Z. Raisi, & J. Zelek, "Investigation of Deep Learning Optimization Algorithms in Scene Text Detection," International Journal of Industrial Electronics Control and Optimization, 2023.
[25] J. P. Norton, "An introduction to identification," Courier Corporation, 2009.
[26] C. J. Miller, S. N. Smith, and M. Pugatch, "Experimental and quasi-experimental designs in implementation research," Psychiatry Research, vol. 283, no. 112452, Jun. 2019.
[27] J. Hamill, G. E. Caldwell, and T. R. Derrick, "Reconstructing Digital Signals Using Shannon's Sampling Theorem," Journal of Applied Biomechanics, vol. 13, no. 2, pp. 226–238, May 1997.

Sadegh Kalantari was born in Tafresh, Iran. He received his B.S. and M.S. degrees in electrical engineering from Tafresh University. He is currently a Ph.D. candidate in control engineering at Tafresh University. His current research interests include system identification, optimization, signal processing, quantum control, quantum computing, geostatistics, and machine learning.

Ali Madadi received a Ph.D. degree in control engineering from the Amirkabir University of Technology. He is an associate professor in the Department of Control Engineering, at Tafresh University. His current research interests include optimal control, robust control, adaptive control, nonlinear control, optimization, and system identification.

Mehdi Ramezani received a Ph.D. in applied mathematics and optimal control from the Amirkabir University of Technology. He is an assistant professor in the Department of Mathematics at Tafresh University. His research interests include optimal control, system identification, optimization, control, soft computing, and machine learning.

Abdolmotaleb Hajati received a Ph.D. degree in mining engineering from Tehran University. He is an assistant professor in the Department of Mining Engineering at Arak University of Technology. His current research interests include nanoflotation and mineral processing.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Multi-Oriented Scene Text Detection at the Character Level

Mahdi Kazeminia[1] | Hamed Shahraki [2] | Mehran Tamjidi [3]

Faculty of Engineering, Velayat University, Iranshahr, Iran.[1, 2, 3]

Corresponding author's email: M.Kazeminia@velayat.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 18-April-2023
Received in revised form: 14-August-2023
Accepted: 16-August-2023
Published online:20-Sep-2023

Keywords:
Deep Learning,
Scene Text Detection,
CRAFT.

ABSTRACT

Recent scene text detection methods perform superior on benchmark datasets using deep-learning frameworks. In this paper, we re-implement the state-of-the-art text detection method, character region awareness for text detection (CRAFT), which can detect individual characters of scene text images. CRAFT is a character-based detection method with many advantages in detecting complex text by detecting character units and estimating the area between characters, capable of detecting texts of any shape. We also improve the detection performance of the baseline method, CRAFT, by some modifications in its architecture and proposing a training scheme that takes benefit of the advanced optimizer. The performance improvements of CRAFT are validated on three benchmark datasets: ICDAR2013, ICDAR2015, and COCO-Text. By applying the pre-trained models on COCO-Text, CRAFT shows that it cannot generalize without fine-tuning. We also improve the ICDAR2015 model and evaluate it on benchmark datasets. The evaluation results show improved precision performance compared to the original pre-trained model with fewer iterations and higher accuracy.

I. Introduction

One of the most popular research topics in computer vision is reading text from wild images [1, 2, and 3], which is the base of numerous practical applications such as multi-language translation, image OCR, and image retrieval. Reading text from scene images is divided into two categories: (1) text detection, which aims at the localization of text in the image, and (2) text recognition, which aims at converting the localized text or cropped word image into a string. This paper focuses on the detection task, which is more challenging than recognition because of the complicated background and significant variance of text shape.

Before the deep learning era, the methods typically identified character or text component candidates using connected component-based or sliding window-based methods, which used hand-crafted features like MSER [4, 5] or SWT [6] as essential components. However, these methods have several significant drawbacks: (1) they are built only to detect individual characters or components, making it challenging for regional context information identification and leading to low recall performance. (2) These methods require multiple post-processing steps for detecting text. (3) They only work on the horizontal type of text and fail on the multi-oriented text.
Therefore, the performance of these classical machine learning-based methods is still far from satisfactory. Recent deep learning-based methods have proved that they can detect more challenging text in scene images. These methods usually adopt general frameworks of object detection methods such as SSD [7], YOLO [8], Faster R-CNN [9], or segmentation frameworks like FCN [10] and Mask R-CNN [11]. Most deep learning-based text detectors that detect text at the word level have difficulty finding curved, extremely long, or highly deformed words with a single bounding box. Character region awareness for text detection (CRAFT) [12] is a character-

based detection method with many advantages in detecting complex text by detecting character units and estimating the area between characters, capable of detecting texts of any shape. Fig. 1 illustrates the CRAFT detection performance for

Fig. 1. The capability of the text detection results in any shape in CRAFT. (a) The polygon bounding box of the detected characters, (b) Gaussian heat-map of individual characters of the text.

different forms of texts. Most existing datasets do not provide character-level annotations and getting character ground truths (GTs) is too expensive. CRAFT localizes the individual character regions and concatenates them into text instances to tackle these problems. Moreover, CRAFT proposes a weakly-supervised learning framework for estimating character-based GTs from word-based annotation datasets to solve the problem relating to the lack of character-based annotations. CRAFT is a segmentation method that outputs a character region score localizing individual characters in an image and a character affinity score grouping characters into a single instance.

In this paper, we improve the detection performance of the baseline method, CRAFT [12], by some modifications in its architecture and proposing a training scheme that takes benefit of the advanced optimizer.

This article is organized as follows: The most significant text detectors are presented in Section II. In Section III, a comprehensive analysis is provided to understand the mechanism of the proposed method. To evaluate the proposed model, some tests are performed in Section IV, and finally, the conclusion is provided in Section V.

II. Background

New deep learning methods see scene text detection as an object detection problem in which words, characters, or text lines are treated as a detection target. For this reason, most text detection methods are based on advanced object detection methods categorized as follows:

F. Regression-Based Text Detectors

These types of methods [13, 14] adapt general Object Detection frameworks such as SSD [7] and Faster R-CNN [9] for text Detection; in these methods, the text is considered as an object, and candidate bounding boxes of text are predicted directly. For example, the single-shot descriptor (SSD) [7] is modified by using long default anchors and filters in TextBoxes [14]. Therefore, it handles the significant variation of aspect ratios of text instances to detect the various type of text shapes. Unlike TextBoxes, the Deep matching prior Network (DMPNet) presented in [15] introduces quadrilateral sliding windows to deal with multi-orientation text. Many regression-based methods [16, 17, and 18] tried to solve the detection challenges of rotated, curved, and arbitrary shape text. For instance, in a recent work, Li et al. [18] proposed a regression-based method that applies only NMS as post-processing to detect arbitrary shape text instances. However, due to structural limitations, it is not very easy to recognize the text of all possible shapes in this way. It would help if we could create anchors that cover all forms and determine the number of proposed boxes, but there is a trade-off between execution time and accuracy.

G. *Segmentation-based Text Detectors*

These methods [19, 20, and 21] classify text regions of images at the pixel level and provide word-level or character-level detection. They usually modified the segmentation framework like FCN [10] and Mask R-CNN [11]. Zhang et al. [22] deployed FCN to predict the salient map of text regions in an oriented scene text. An attention procedure is used in a single-shot text detector (SSTD) [19] to boost the text regions of an image and reduce the effect of background interference on the feature level. FCN is also deployed in TextSnake [21] as a base detector. The aim is to extract text instances by detecting and assembling local components. Liao et al. [23] proposed a simple segmentation network that utilizes differentiable binarization in recent work. It can real-time detect text in arbitrary shapes.

H. *End-to-End Text Detectors*

In these methods [24, 25] the detection and recognition mechanisms are trained simultaneously. End-to-end text detector methods improve the performance through the recognition module. FOTS [24] proposed the RoIRotate to effectively detect horizontal and multi-oriented text of an image by sharing convolutional features between detection and recognition modules. In [25], a single-shot text spotter is presented to make a direct mapping between an input image

and character sequences by learning. The performance of detection is high in end-to-end methods because they deploy the recognition results to improve the text detection performance. In [26], an end-to-end segmentation-based method is presented to predict a character-level probability map. It uses the Mask R-CNN architecture [11] for detecting and recognizing text instances in arbitrary shapes.

The previously mentioned methods aim to detect words in the input image. However, it is challenging to use a word as the basic unit of detection, leading to some problems. These problems are: (1) It is difficult to cope with arbitrary shape text: In previous methods, the training process was performed with rigid word-level bounding boxes, and limitations are exhibited to represent the text region in an arbitrary shape [27], (2) Defining the word is not easy: Although, in most methods, the text is detected with words as its unit, determining the extent of a word is essential for detection. Because words are distinguished by different bases, such as spaces or color, some letters do not have spaces, (3) it is evident that character is the basic unit in handling various text types: scene text images are organized hierarchically for visual components, such as characters, text lines, and words. As most languages are established based on characters, performing text detection methods on character detectors seems reasonable [28].

I. *Character-Level Text Detectors*

Zhang et al. [29] utilized MSER [4] to detect text at the character level. However, this method performs poorly on the low contrast images and curve shape texts. In [30], text maps obtained by segmentation use character-level annotations to generate multi-oriented text bounding boxes. Seglink [31] searches for small text elements (segments) in the image and links these segments to create word boxes with additional post-processing. CRAFT improved the idea of WordSup [28] by using a weakly supervised framework to detect individual characters in arbitrary shape text, enabling it to achieve state-of-the-art performance on public benchmark datasets.

III. Methodology

A. Architecture

Fig. 2 shows the network architecture of CRAFT. The overall architecture is similar to U-Net [32], a standard segmentation model based on FCN [10]. It adopts VGG-16 [33] as a backbone in the encoder and generates a Region Score/affinity score map by applying up-sampling and skip-connection of the U-Net structure. In simple words, VGG16 is essentially the feature-extracting architecture that is used to encode the network's input into a certain feature representation. The decoding segment of the CRAFT has to skip connections that aggregate low-level features. It predicts two scores for each character:

Region Score: As the name suggests, it gives the region of the character. It localizes the character.

Affinity Score: 'Affinity' is the degree to which a substance tends to combine with another. So, an affinity score merges characters into a single instance (a word). CRAFT generates

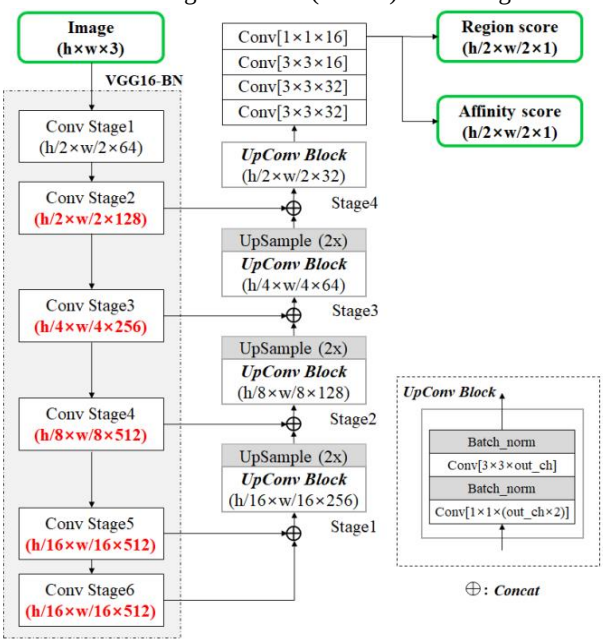


Fig. 2. The network architecture of CRAFT [1]. The size of the feature layers has been miscalculated, which is corrected in this figure with red.

two maps as output: Region Level Map and Affinity Map. The areas where the characters are present are marked in the Region Map. The Affinity Map is a pictorially that represents the related character. Finally, the affinity and region scores are combined to give the bounding box of each word. The coordinates are in order:

(left-top), (right-top) (right-bottom), (left-bottom), where each coordinate is an (x, y) pair.

B. Training

Ground Truth Label Generation**:** A GT label is created for each learning image's region and affinity score with character bounding boxes. Region scores and affinity scores are encoded as Gaussian heat-map. The heat map representation is highly flexible when dealing with GT without strict limitations. The proposed GT definition makes detecting large and long text possible even if the model's receptive fields are small. Character-by-character detection allows the convolution filter to focus on the character-to-character relationship instead of the entire text instance.

Weakly-Supervised Learning: Unlike Synthetic image datasets, real datasets usually provide only word-by-word annotations. CRAFT creates a text box from a word-by-word annotation weakly supervised. The trained model generates a character bounding box by predicting a region score from an

TABLE I
TEXT DETECTION DATASETS USED IN THE CRAFT[12]

Dataset	Language	Year	# Images			# Text instance			Text Shape		Annotation level		
			Total	Train	Test	Total	Train	Test	AQ	MO	Char	Word	Text-Line
MSRA-TD500 [34]	EN/CN	2012	500	300	200	-	-	-	✓	–	–	–	✓
ICDAR2013 [35]	EN	2013	462	229	233	1944	849	1095	–	–	✓	✓	–
ICDAR2015 [36]	EN	2015	1500	1000	500	17548	122318	5230	✓	–	–	✓	–
SynthText [37]	EN	2016	800k	-	-	6M	-	-	✓	–	–	✓	–
Total-Text [38]	EN	2017	1525	1225	300	9330	-	-	✓	✓	–	✓	✓
CTW-1500 [39]	EN/CN	2017	1500	1000	500	10751	-	-	✓	✓	–	✓	✓
MLT-2017 [40]	ML	2017	18000	7200	10800	-	-	-	✓	–	–	✓	–
COCO-Text [41]	EN	2014	63686	43686	20000	145859	118309	27550	✓	–	–	✓	–

Note: H: Horizontal, MO: Multi-Oriented, C: Curved, EN: English, CN: Chinese, ML: Multi-Language, D: Detection, R: Recognition, AQ: Arbitrary-Quadrilateral.

image in which a word unit annotation image is provided. To show the reliability of the intermediate model prediction, the reliability map for each word box is calculated in proportion to the number of detected characters divided by the number of characters in the GT. This is used as a learning weight learning. The real datasets can be trained as follows:

- According to the label of the real-world data, the text line is cropped out.
- Run the network and get the resulting map.
- Split a single text based on the graph output from the network.
- Generate a label based on the results of the previous step.

These labels are not necessarily accurate, so training with these data does not necessarily make the model better, especially when the model is not accurate at first. Therefore, the paper uses an existing annotation result to give confidence that the text length is used. If the length of the text is the same as the result of the model, then the confidence is high. Otherwise, the confidence is low (the confidence is directly multiplied by the loss).

For generating pseudo-GT, the confidence is calculated from the length information of the text to determine whether the data is reliable, and the confidence is used to update the loss. The confidence of each word is calculated as follows:

$$S_{conf}(w) = \frac{l(w) - \min\left(l(w), |l(w) - l^c(w)|\right)}{l(w)} \quad (1)$$

Where l(w): word length of the given word-level annotation. $l^c(w)$: is the length of the word (number of bounding boxes) obtained from the score map. $0 < S_{conf}(w) < 1$: when the actual number of boxes equals the estimated number of boxes, close to 0 when different.

$$S_c(p) = \begin{cases} S_{conf}(w) & p \in R(w) \\ 1 & otherwise \end{cases} \quad (2)$$

Where, p: each pixel, R(w): area occupied by word w, Sc(p): is a confidence value for each pixel, and the Sc map has a confidence value of 0 to 1 on a pixel-by-word basis and the loss function is computed as follow:

$$L = \sum_p S_c(p)\left(\|S_r(p) - S_r^*(p)\|_2^2 + \|S_a(p) - S_a^*(p)\|_2^2\right) \quad (3)$$

Where, $S_r^*(p)$ and $S_a^*(p)$ are pseudo-ground truth region scores and affinity map, respectively. $S_r(p)$ and $S_a(p)$ are predicted region scores and affinity scores. They are used to localize the individual characters in the image, and group each character into a text instance, respectively. As learning progresses, the CRAFT model predicts letters more accurately, and the S_{conf} also gradually increases.

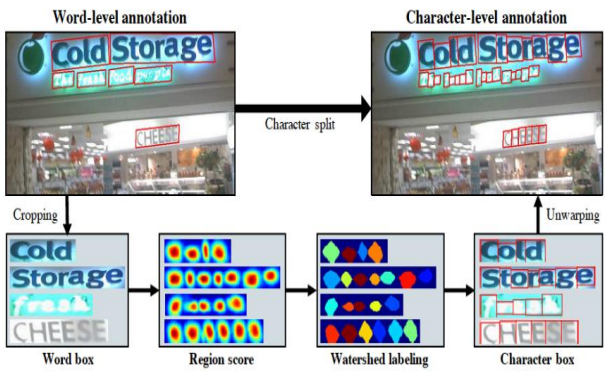


Fig. 3. Character-level annotation procedure of CRAFT [12].

C. Post Processing

In the post-processing stage, the output of the final stage can be varied in terms of word boxes, text boxes, and polygons.

After the network outputs the score map, the pixel-level labels are combined into a box. For this purpose, first, the threshold map is used to filter the score map, and a connected domain analysis is performed (connected component labeling), then draw the final quadrilateral bounding box through the connected domain.

D. Improvement

CRAFT has two significant drawbacks. Firstly, it is slow during training, especially for fine-tuning real-word datasets producing many post-processing stages for character-level annotation and using these labels for training. Secondly, the recall performance of detection does not improve fast. To address the first issue, we change the learning rate hyperparameters. To address the second issue, we modify the Gaussian heat map and watershed labeling thresholds in the

post-processing framework of Fig. 3.

IV. Experiments

A. Datasets

Table I shows the benchmark datasets used in CRAFT [12]. We trained our model on ICDAR2015 and COCO-Text and evaluated ICDAR13, ICDAR15, and COCO-Text.

1) ICDAR2013: This dataset was introduced in the ICDAR 2013 Robust Reading Competition for focused scene text detection [35]. This dataset is annotated at the word level using rectangular boxes containing 229 and 233 images for training and testing, respectively.

2) ICDAR2015: This dataset was introduced in the ICDAR2015 for incidental scene text detection, which contains 1000 images for training and 500 image testing. The annotations of this dataset are at the word level using quadrilateral boxes.

3) ICDAR2017: This dataset is one of the enormous multi-lingual text datasets, including scene images in nine languages. ICDAR17 Multi-language dataset has 18000 images with 7200 training images, 1800 validation images, and 9000 testing images [40].

4) COCO-Text: This dataset introduced in [41] is the largest and most challenging text detection dataset, consisting of 43,686 annotated images used for the training and 10,000 images for training [42, 30]. The images are annotated with rectangle bounding boxes in this dataset at the word level.

5) Synth Text: In the Wild dataset [43], the SynthText contains 858,750 synthetic scene images with 7,266,866 word-instances and 28,971,487 characters. Moreover, most text instances in the data set have multi-oriented and annotated features with word and character-level rotated bounding boxes and text sequences. These instances are created by blending natural images with text rendered with random fonts, sizes, orientations, and colors.

B. Training Procedure

In summary, the training procedure is (1) an image-net pre-trained model, (2) pre-train on SynthText, and (3) fie-tune on real datasets. We perform all of our experiments on two NVIDIA Tesla T4 GPUs with 8GB RAM. We use Adam [44] as an optimizer and set the initial learning rate to 3.278e - 5. We multiply the learning rate by 0.8 at 5k iterations. We train our model for about three epochs, use a batch size of 8, and resize the input images to 768 × 768. We re-implemented the training code of CRAFT, and our models are trained from scratch. First, we trained CRAFT on the Synthetic dataset for 50k iterations and weakly supervised the COCO-Text and ICDAR2015 datasets for 75 and 175 epochs, respectively. For SynthText and COCO-Text, we follow the [12] when settings the hyperparameters, but for the ICDAR2015 model, we modified the learning rate and the post-processing threshold parameters to improve the performance of the paper.

C. Experimental Results

This dataset is one of the enormous multi-lingual text datasets, including scene images in nine languages. ICDAR17 Multi-language dataset has 18000 images with 7200 training images, 1800 validation images, and 9000 testing images [40].

$$H - mean = 2\frac{P \times R}{P + R} \quad (4)$$

We compare the effects of the ICDAR2013 and ICDAR2015 datasets that have been used in the [12]. Besides, a new dataset, namely COCO-Text, is used for evaluation, which has not been used in the main paper. Table II compares the IOU results of the pre-trained models.

Pre-trained models of the paper: In the official Github1 page of the CRAFT [12], researchers provide two models; The first model, Git_Model_MLT, is fine-tuned on a combination of ICDAR2013 [35] and ICDAR17-MLT [40] training images, and the model Git_Model_IC15 is fine-tuned only on the ICDAR2015 [36] dataset. Using the model Git_Model_IC15, H-mean declined by about 5% for ICDAR2015 and COCO-Text compared to model Git Model MLT and about 15% for the ICDAR2013 dataset. Thus, CRAFT requires more images for training to achieve better results. Therefore, to tackle this problem, we can use more data for training the craft. Fig. 4 compares the output results of two models on ICDAR2013 and ICDAR2015. As seen in Table II, CRAFT's performance for the two models declined significantly on the COCO-Text dataset. These models do not have generalization ability without fine-tuning on unseen datasets. Fig. 5 shows the output results of the two models on the COCO-Text.

Our trained models: As shown in Table II, we provide three models for CRAFT that are trained from scratch. The first model, Syndata, is trained on 800k train images of the SynthText [37] dataset, and for the training of this model, we follow the same structure and hyperparameters used in the CRAFT. Although this dataset has many images for training, this model performs poorly on real datasets because the text in real datasets is different from synthetic datasets.

For creating the Fine_Coco model, we fine-tuned the pre-trained Syndata model on the COCO-text dataset with many image and text instances for training. This model shows excellent precision performance in evaluation, especially in the COCO-text datasets compared to the Git_Model_IC15 of CRAFT. Still, the recall performance of this model was low because this model only trained for 75 epochs and needs optimization of hyperparameters.

The major limitation of the CRAFT framework is that it is too slow during fine-tuning of real datasets because it has much post-processing for the detection of individual characters in the image and generating word bounding boxes. For example, the authors of CRAFT used 4 GPU and 500 epochs to fine-tune the Git_Model_IC15 model. However, we fine-tuned the Fine_IC15 model for 175 epochs and two GPUs and achieved better results in comparison to the model of CRAFT (Table II).

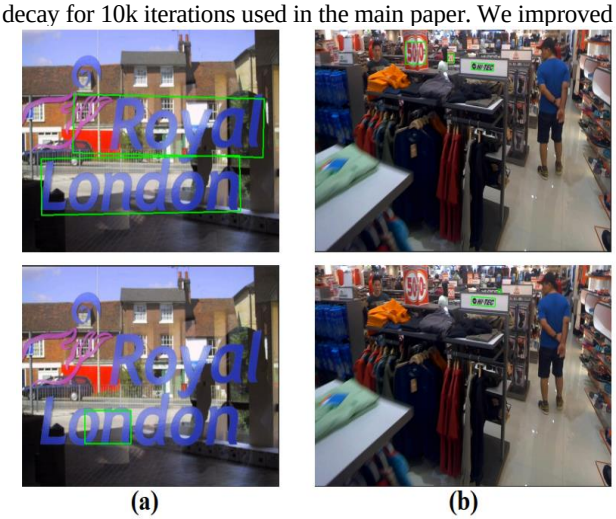
TABLE II

THE BASELINE AND OUR TRAINED MODEL ARE COMPARED USING IOU [36] EVALUATION METRIC. THE BEST PERFORMANCES ARE SHOWN IN BOLD. PROPOSED MODELS PERFORM BETTER THAN THE BASELINE MODELS (AUTHORS) FOR THE DATASETS IN THE STUDY.

	Model	ICDAR2013*			ICDAR2015			COCO-Text		
		Precision	Recall	H-mean	Precision	Recall	H-mean	Precision	Recall	H-mean
Authors	Paper	97.4	93.1	95.2	89.8	84.3	86.9	–	–	–
	Git_Model_MLT	92.2	90.4	91.3	88.9	80	84.2	64.6	58.2	61.2
	Git_Model_IC15	72.7	77.6	75.2	82.2	77.8	79.9	56.7	55.9	56.3
Ours	Syndata	75.8	66.8	71	70.8	38.5	49.9	46.8	22.5	30.4
	Fine_Coco	80.8	**77.5**	79.2	82.1	62.2	70.8	**70.0**	50.8	**58.9**
	Fine_IC15	**83.7**	76.7	**80.1**	**85.5**	**76.3**	**80.6**	62.0	**53.9**	57.7

*: This dataset is evaluated on the DetEval [35] evaluation metric in the paper, while here I used the IOU [36] metric.

For this purpose, we used a learning rate of 3.2768e-5 and decay of 0.8 for every 5k iterations compared to 1e-4 and 0.8 decay for 10k iterations used in the main paper. We improved precision performance by about 11%, 3%, and 4% in the ICDAR2013, ICDAR2015, and COCO-Text datasets.



(a) (b)

Fig. 4. The detected text of model Git_Model_MLT (above) and model Git_Model_IC15 (below) on (a) ICDAR2013, and (b) ICDAR2015 datasets. The model Git_Model_MLT trained on more images performs better than the model Git_Model_IC15.

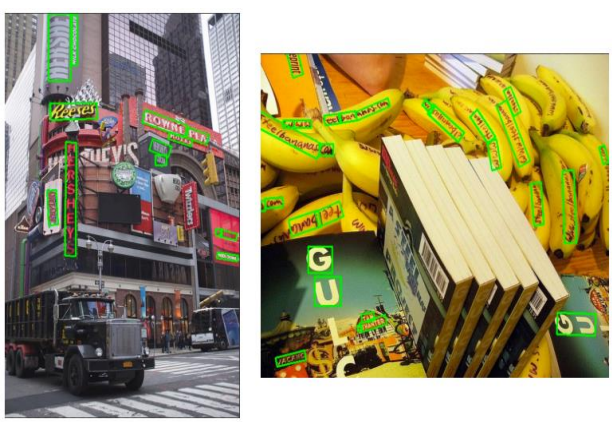
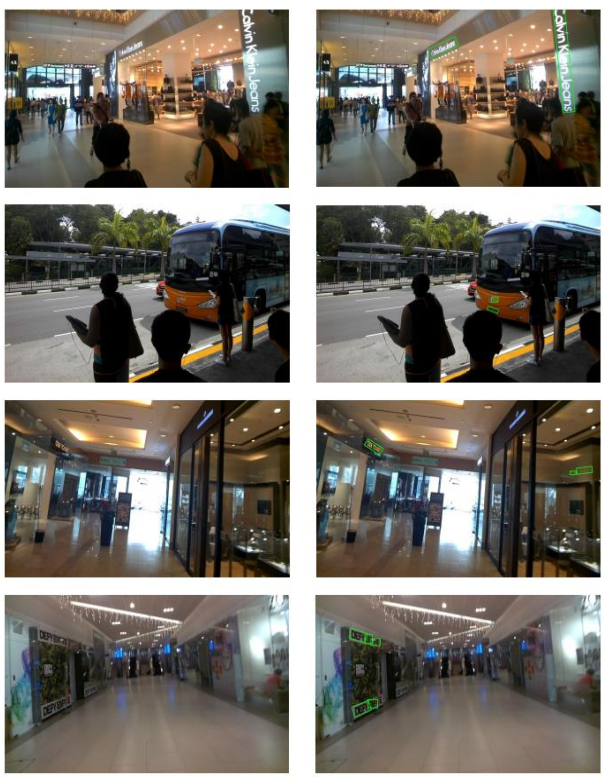


Fig. 5. The output result of our final model on the COCO-Text dataset. The baseline CRAFT model performed poorly on the occluded and low-resolution text of these images. These results are obtained from the best model of CRAFT (Git Model MLT).



(a) (b)

Fig. 6. Comparison of some selected images of the ICDAR2015 dataset: (a) The detection result of the Git_Model_IC15 model from the main paper, which failed to detect any text in these images, and (b) the detection results of our trained model, which performed better.

We also trained CRAFT on the COCO-Text dataset, in which we achieved a 14% precision improvement compared to the Git_Model_IC15 of the CRAFT. Fig. 6(a) shows some selected images from the ICDAR2015 dataset in which the pre-trained model (Git_Model_IC15) failed to detect text in them, and Fig. 6(b) shows the detection result of our trained model that performs better in these images.

We also compare our model with several previous [12, 13,

TABLE III
COMPARISON OF OUR MODEL WITH SEVERAL WELL-KNOWN PREVIOUS AND RECENT STATE-OF-THE-ART TEXT DETECTION METHODS.

Method	ICDAR13			ICDAR15			COCOText		
	Precision	Recall	H-mean	Precision	Recall	H-mean	Precision	Recall	H-mean
EAST [13]	**84.8**	74.24	79.2	83.6	73.5	78.2	55.4	32.8	41.3
Pixellink [20]	62.2	62.5	62.3	82.8	81.6	82.2	61.0	33.4	43.2
PSENET [45]	81.0	62.4	70.5	84.6	77.5	80.9	60.5	49.3	54.4
DBNet [23]	--	--	--	86.8	78.4	82.3	--	--	--
DRRG [46]	--	--	--	88.5	84.7	86.6	--	--	--
TextDCT[47]	--	--	--	**88.9**	**84.8**	**86.8**	--	--	--
Baseline [12]	72.7	77.6	75.1	82.2	77.8	82.2	56.7	**55.9**	56.3
Our_Model	83.7	**76.7**	**80.1**	85.5	76.3	80.6	**62.0**	53.9	**57.7**

20, 45] and recent state-of-the-art [23, 46, 47] scene text detection methods in Table III on three benchmark datasets [35, 36, and 41].

For a fair comparison, we follow a similar procedure described in [48] for comparing our model with the previous methods in [12, 13, 20, 45]. Table III shows that our model outperformed these approaches regarding H-mean performance in all three benchmarks. Compared to the baseline [12] model with a similar setup of training and parameters, our model also obtained a higher margin of H-mean performance. These performances are evident in ICDAR13 (~5% higher) and ICDAR15 (~3% higher). For the unseen dataset, COCO-Text also achieved the best results in terms of precision and H-mean compared to the baseline model.

To further enhance our model's performance in scent text detection, we plan to explore various cutting-edge techniques in natural language processing and computer vision. This includes investigating the latest advancements in deep learning architectures, such as transformer-based models and attention mechanisms, which have shown great promise in tackling complex language tasks. Additionally, we will focus on incorporating domain-specific knowledge and contextual information to better understand scent-related text patterns and improve the model's ability to discern relevant information. Furthermore, data augmentation techniques will be utilized to create a more diverse and comprehensive training dataset, enabling our model to generalize better to real-world scenarios. By continuously refining and optimizing the model's architecture, we are confident in bridging the performance gap and establishing our model as a competitive solution for scent text detection tasks.

V. Conclusions

In this paper, we have re-implemented CRAFT. This text detection method uses a weakly supervised learning method to detect the character of scene text images CRAFT achieved state-of-the-art performance in most public benchmark datasets. By applying the pre-trained model provided by CRAFT researchers, we showed that CRAFT does not have the generalization capability on unseen datasets without fine-tuning. We also made some modifications to the model, which was fine-tuned on the ICDAR2015 dataset and improved CRAFT's performance with fewer iterations and higher accuracy compared to the model provided by the authors.

References

[1] H. Lin, P. Yang, and F. Zhang, "Review of scene text detection and recognition," *Archives of Computational Methods in Engineering*, pp. 433–454, 2020.

[2] S. Long, Y. Guan, B. Wang, K. Bian, and C. Yao, "Rethinking Irregular Scene Text Recognition," *arXiv.org*, Nov. 11, 2019. https://arxiv.org/abs/1908.11834.

[3] T. Diep, "State-of-the-art in action: unconstrained text detection, " in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0-0.

[4] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761–767, Sep, 2004.

[5] L. Neumann and J. Matas, "A method for text localization and recognition in real-world images," in *Asian Conference On Computer Vision, Springer*, 2010, pp. 770-783.

[6] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform, " in *Proc. IEEE Conf. on Comp. Vision and Pattern Recognit*, 2010, pp. 2963–2970.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and A. Berg, "SSD: single shot multibox detector, " in *European Conference on Computer Vision, Springer*, Oct. 2016, pp. 21–37.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. "You only look once: Unified, real-time object detection, " in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.

[9] S. Ren, K. He, R. Girshick, and J. Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing*

Systems, 2015, pp. 91–99.
[10] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation, " in *Proceedings Of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
[11] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proc. Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2961–2969.
[12] Y. Baek, et al. "Character region awareness for text detection" in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9365-9374.
[13] X. Zhou, C. Yao, H. Wen, Y. Wang, S. Zhou, W. He, and J. Liang, "East: an efficient and accurate scene text detector," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5551–5560.
[14] M. Liao, B. Shi, X. Bai, X. Wang, and W. Liu, "Textboxes: A fast text detector with a single deep neural network," in *Thirty-First AAAI Conference on Artificial Intelligence*, Feb. 2017.
[15] Y. Liu and L. Jin, "Deep matching prior network: Toward tighter multi-oriented text detection, " in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1962–1969.
[16] M. Liao, B. Shi, and X. Bai, "Textboxes++: a single-shot oriented scene text detector," *IEEE Transactions on Image Processing*, Vol. 27, No. 8, pp. 3676–3690, Apr. 2018.
[17] J. Ma, W. Shao, H. Ye, L. Wang, H. Wang, Y. Zheng, and X. Xue, "Arbitrary-oriented scene text detection via rotation proposals, " *IEEE Transactions on Multimedia*, Vol. 20, No. 11, pp.3111–3122, 2018.
[18] X. Li, J. Liu, S. Zhang, and G. Zhang, "Learning to predict more accurate text instances for scene text detection," *arXiv preprint arXiv: 1911.07423*, 2019.
[19] P. He, W. Huang, T. He, Q. Zhu, Y. Qiao, and X. Li, "Single shot text detector with regional attention," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 3047–3055.
[20] D. Deng, H. Liu, X. Li, and D. Cai, "Pixellink: Detecting scene text via instance segmentation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32, No. 1. 2018.
[21] S. Long, et al. "Textsnake: A flexible representation for detecting text of arbitrary shapes," In *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 20–36.
[22] Z. Zhang, C. Zhang, W. Shen, C. Yao, W. Liu, and X. Bai, "Multi-oriented text detection with fully convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4159–4167.
[23] M. Liao, Z. Wan, C. Yao, K. Chen, and X. Bai, "Real-time scene text detection with differentiable binarization," In *AAAI Conf. on Artificial Intelligence*, 2020, pages 11474–11481.
[24] X. Liu, D. Liang, S. Yan, D. Chen, Y. Qiao, and J. Yan, "Fots: Fast oriented text spotting with a unified network," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5676–5685.
[25] T. He, Z. Tian, W. Huang, C. Shen, Y. Qiao, and C. Sun, "An end-to-end text spotter with explicit alignment and attention," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5020–5029.
[26] P. Lyu, M. Liao, C. Yao, W. Wu, and X. Bai, "Mask textspotter: An end-to-end trainable neural network for spotting text with arbitrary shapes," *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 67-83.
[27] Q. Wang, Y. Zheng, and M. Betke, "Sa-text: Simple but accurate detector for text of arbitrary shapes," arXiv preprint arXiv:1911.07046, 2019.
[28] H. Hu, C. Zhang, Y. Luo, Y. Wang, J. Han, and E. Ding, "Wordsup: Exploiting word annotations for character based text detection," *in Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4940–4949.
[29] S. Zhang, M. Lin, T. Chen, L. Jin, and L. Lin, "Character proposal network for robust text extraction," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 2633–2637.
[30] C. Yao, X. Bai, N. Sang, X. Zhou, S. Zhou, and Z. Cao, "Scene text detection via holistic, multi-channel prediction," arXiv preprint arXiv: 1606.09002, 2016.
[31] B. Shi, X. Bai, and S. Belongie, "Detecting oriented text in natural images by linking segments," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2550–2558.
[32] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.
[33] K. Simonyan and A. Zisserman, "Detecting oriented text in natural images by linking segments," CoRR, abs/1409.1556, 2014.
[34] C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, "Detecting texts of arbitrary orientations in natural images," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2012, pp. 1083–1090.
[35] D. Karatzas, et al. " Icdar 2013 robust reading competition, " in *12th international conference on document analysis and recognition*, 2013, pp. 1484–1493.
[36] D. Karatzas, et al. "Icdar 2015 competition on robust reading," in *13th international conference on document analysis and recognition (ICDAR, 2015)*, pp. 1156–1160.
[37] A. Gupta, A. Vedaldi, and A. Zisserman, "Synthetic data for text localisation in natural images, " In *Proc. IEEE Conf. on Comp. Vision and Pattern Recognit*, 2016, pp. 2315–2324.
[38] C.K. Ch'ng and C.S. Chan, "Total-text: A comprehensive dataset for scene text detection and recognition," In *Proc. IAPR Int. Conf. on Document Anal. and Recognit.* (ICDAR), Vol. 1, 2017, pp. 935–942.
[39] L. Yuliang, J. Lianwen, Z. Shuaitao, and Z. Sheng, "Detecting curve text in the wild: New dataset and new solution," In arXiv preprint arXiv: 1712.02170, 2017.
[40] M. Iwamura, N. Morimoto, K. Tainaka, D. Bazazian, L. Gomez, and D. Karatzas, "Icdar2017 robust reading challenge on omnidirectional video," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Vol. 1, 2017, pp. 1448–1453.
[41] T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision, Springer*, 2014, pp. 740–755.
[42] A. Veit, T. Matera, L. Neumann, J. Matas, and S. Belongie, "Coco-text: Dataset and benchmark for text detection and recognition in natural images," arXiv preprint arXiv:1601.07140, 2016.

[43] Z. Raisi, M.A. Naiel, P. Fieguth, S. Wardell, and John Zelek. "Text detection and recognition in the wild: A review, " arXiv preprint arXiv: 2006.04305, 2020.
[44] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization, " arXiv preprint arXiv:1412.6980, 2014.
[45] W. Wang, E. Xie, X. Li, W. Hou, T. Lu, G. Yu, and S. Shao, "Shape robust text detection with progressive scale expansion network," arXiv preprint arXiv:1903.12473, 2019.
[46] S.X. Zhang, et al. "Deep relational reasoning graph network for arbitrary shape text detection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 9699–9708.
[47] Y. Su, Z. Shao, Y. Zhou, F. Meng, H. Zhu, B. Liu, and R. Yao, "Textdct: Arbitrary-shaped text detection via discrete cosine transform mask", *IEEE Transactions on Multimedia*, 2022.
[48] Z. Raisi, et al. "Smart Text Reader System for People who are Blind Using Machine and Deep Learning," *Machine Learning Algorithms for Signal and Image Processing*, pp. 161-200, 2022.

Mahdi Kazeminia received the B.Sc., the M.Sc., and his Ph.D degree in telecommunication engineering from University of Sistan and Baluchestan, Iran, Zahedan in 2010, 2012, and 2019, respectively.
From 2017 to 2018, he was visiting researcher at the Univesity of Padova, Padova, Italy. Currently, he is an assistant professor in the Department of Electronics, Velayat University, Iranshahr, Iran. His research interests include resource allocation optimization, IoT networks, and D2D communication.

Hamed Shahraki was born in Zabol, Iran in 1984. He received BSc in Electrical Engineering from University of Sistan and Baluchestan, Zahedan, Iran in 2008, MS in Electrical Engineering from Shiraz University of Technology, Shiraz, Iran in 2012, and Ph.D. degree from Shahid Bahonar University of Kerman, Kerman, Iran in 2018. He is currently an Assistant Professor with the Electrical Engineering Department, Velayat University of Iranshahr, Iranshahr, Iran. His main areas of research interest are MTM and RF/micro wave circuits design.

Mehran Tamjidi is currently an Assistant Professor in the Department of Industrial Engineering at the University of Velayat. He received a B.S. in Industrial Engineering from the Azad University, Zahedan branch in 2005, and a M.S. in Manufacturing Engineering from University Malaya (UM) in 2012, and continued on to receive the Ph.D. in Manufacturing Systems Engineering from University Putra Malaysia (UPM) in 2017. His research interests include the machining process and optimization technics.

IECO

This page intentionally left blank.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Risk-Cost Minimization in Optimal Reactive Power Dispatch Problem in the DFIG Integrated System

Meysam Mokari[1] | Mohammad Hasan Moradi[2] | Mohammad Abedini[3]
Iran Grid Management Company, Tehran, Iran.[1]
Faculty of Electrical and computer Engineering, University of Buali-Sina, Hamedan, Iran.[2]
Faculty of Electrical Engineering, University of Ayatollah Bourojerdi, Bourojerd, Iran.[3]
Corresponding author's email: mhmoradi@basu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 07-July-2023
Received in revised form: 22-August-2023
Accepted: 24- September-2023
Published online: 24- September-2023

Keywords:
Multi-objective risk-based,
Optimal reactive power dispatch,
Voltage instability Risk,
Power system uncertainty,
Hybrid multi objective PSO.

ABSTRACT

Objective: In this paper, a novel method for a multi-objective and risk-based optimal reactive power dispatch is proposed. The method includes two main objective functions: technical and economic. The technical objective involves minimizing the risks of voltage instability, voltage deviation, and flow violation, and the economic objective involves minimizing the costs of reactive power generation, active power losses, load shedding, and active power rescheduling. Using these functions and assigning different weighting factors for each sub-objective, the risk of the events or uncertainties to customers or the grid can be managed. In addition, moment matching is used to discretize and create scenarios from continuous probability distribution functions of wind speed and electrical energy uncertainties. As the number of uncertain variables increases, so does the number of scenarios and the simulation time. Therefore, the fast-forward selection algorithm is applied to reduce the number of scenarios. To reduce the computational complexity and the number of topological scenarios, a new contingency filtering method based on high-risk events is proposed. A modified multi-objective PSO algorithm based on a hybrid PSO with sine-cosine acceleration coefficients is proposed to find the Pareto front of solutions. The method is implemented on the modified IEEE 30-bus test system. To demonstrate the effectiveness of the proposed method, the results are compared with previously published literature. The results show that risk-based scheduling increases system reliability and cost-effectiveness compared to traditional scheduling.

NOMENCLATURE

fr	Failure rate of the component per year
rr	Repair rate of the component per year
k	Shape factor of Weibull distribution
c	Scaling parameter of Weibull distribution
v	Wind speed (m/sec)
$a_{gi}^{q}, b_{g}^{q}, c_{gi}^{q}$	Reactive power cost coefficient
$g_{k,\ B_k}$	Conductance and suseptance of the line
i,j	Bus i and event j
S	Scenario number
EV, ER	Emergency voltage limit, Emergency rating of the line
VIR, VVR, FVR	Voltage instability risk,Voltage violation risk, flow violation risk
VSM	Voltage stability margin
V_{min}	Minimum voltage of the network
SL_i	Apparent power of line i
P	Average active power price in $/MW
PF	Participation factor of the generators
$NC, n_{stage}, n_{lc-load}$	Number of contingencies, number of SPS stages, number of load cut performed
OFV, IF	Objective function value, importance factors of the sub-objective functions
N.OUT	Number of outputs

I.Introduction

A. Research Motivation

Optimal Reactive Power Dispatch (ORPD) is a special form of Optimal Power Flow (OPF) that deals with reactive power instead of active power. The objective of this optimization problem is to find the optimal value of reactive power resources in the system to improve security and economic efficiency. Hence, it is one of the most principal problems in power systems. However, system security is usually compromised after an event.

To ensure system security during events, it is essential to simulate contingencies in the planning phase (a preventive or corrective approach). When contingencies are simulated, the problem is treated as a security constraint OPRD (SC-ORPD). The SC-ORPD, which is a deterministic method, has been used by independent system operators (ISO) for several years because of its high reliability and simplicity. Operational and security constraints must be met before and after contingencies. However, deterministic approaches are often too conservative and therefore not economical (especially in preventive planning), since they are based on the most severe events [1]. Therefore, the use of these methods is less and less considered by operators and researchers nowadays.

In addition, the installation of renewable energy sources (RESs) in power systems has recently increased due to environmental issues. It has been shown that the deterministic modelling of renewables can lead to voltage instability and power flow fluctuations due to their intermittent nature. Demand uncertainty due to load forecasting errors is also unavoidable. Demand uncertainty also affects power system planning, security, and economics. Therefore, these uncertainty parameters must be efficiently considered in the planning and operation phases to ensure system security and economics.

Recently, the risk-based ORPD (RB-ORPD) has received more attention because of its ability to address the above issues. The risk indices quantify the system stress and can be formulated using the probability and consequence of the events or the uncertainties (usually the production of probability and consequence). These functions can be calculated based on the impact of the events on the customers or the transmission system.

It is worth noting that risk-based methods are so efficient that can not only increase system safety (by taking into account contingencies or uncertainties of the system in problem formulation) but also increase system efficiency (due to allocation of resources to high consequence and probability events rather than only to high consequence ones).

However, reducing system risk leads to an increase in operating costs (these are conflicting goals). Therefore, for ORPD or OPF problems, a comprehensive techno-economic objective function must be considered to ensure the reduction of system or customer risk and operating costs. The objective function should include all or some of the various system risks (risk of voltage instability, risk of voltage violation, etc.), customer risks (loss of load expectancy, loss of load frequency, etc.), and operating costs (cost or price of active or reactive power generation, system losses, etc.).

Furthermore, if scheduling is expected to be implemented with minimal deviation in the control centers of the power system, the practical limitations of the sources and control systems must be included in the formulation of the ORPD problem. Thus, the problem must be subjected to various practical and mathematical constraints.

Considering the previously mentioned objectives and constraints, the ORPD is difficult to solve. From the point of view of solving the optimization problem, the ORPD is a nonlinear and nonconvex problem with many constraints. Many traditional optimization methods have been presented in the literature to solve this problem. However, these methods are trapped in local optima, and some depend on the initial estimation of the variables and do not provide much freedom in choosing the objective functions and the different constraints. Therefore, metaheuristic methods and evolutionary algorithms can be used in this area.

B. Literature Review

The ORPD problem has been addressed in several research studies. In [2], a new corrective voltage control (CVC) for power systems is presented. The uncertainty of wind power generation and demand values is considered in the problem formulation through a scenario-based approach. Moreover, the participation of demand-side resources is considered an effective control option to reduce the control cost. [3] presents a novel stochastic preventive voltage control, ORPD. Wind speed uncertainty is considered in this reference and a scenario-based modelling approach is used to discretize the continuous PDF of uncertain variables. The risk associated with each objective function is calculated using the conditional value at risk (CVaR). In [4], a new criterion for real-time estimation of "N-1" system risk is presented. The risk is calculated using the product of probability and impact. In [5], a probabilistic optimal power flow considering wind uncertainty is presented. The objectives of the presented method are to minimize the expected generation cost and the downside risk. In [6], a framework with weighted random constraints based on optimal power flow is presented. The uncertainty of the wind power forecast is taken into account in this reference.

In [7], the authors propose a new framework for stochastic OPF. The framework can balance the generation cost and system risk. In [8], a novel scenario-based method for avoiding voltage instability under wind and load uncertainty is presented. In [9], the optimal placement and sizing of the static var compensator (SVS) are presented to minimize the risk of short-term voltage stability after the occurrence of events. The objective functions are to minimize the investment cost of

SVC and the risk of short-term voltage instability. The problem is solved using a genetic algorithm. The paper considers the effects of DGs and the customers' dynamic load model. However, the effects of the uncertainties of DG are neglected.

In [10], coordinated optimization of reactive power reserve, stochastic and multi-objective optimal power, and reactive power management are presented. Economic objectives and an increase in the load margin are considered objective functions to minimize. In [11], a techno-economic algorithm for SC-ORPD is presented, which considers the uncertainty of wind and demand. MCS is used to incorporate stochastic parameters, and NSGA-II is used to find non-dominant solutions. In [12], a method for calculating probability densities of the voltage level is presented. For this purpose, the problem is integrated with an ORPD that takes into account the uncertainty of active power generation. In this work, the probability densities are handled with MCS and used as soft constraints for an ORPD problem. In [13], the ORPD problem with minimizing the loss, voltage deviation, and increasing the voltage stability margin is presented. The problem is formulated as a mixed-integer nonlinear optimization problem that contains both discrete and continuous decision variables. The Artificial Bee Colony (ABC) algorithm is used to solve the problem.

In [14], a risk-based distribution robust optimization problem is proposed to balance the operational cost and risk in the real-time framework. In this reference, the presented method takes into account the voltage reliability of the buses even if the PDF of the uncertain variables cannot be estimated accurately. [15] The Rao-3 optimization algorithm is used to solve the constrained ORPD problem. In addition, the uncertainties of solar and wind energy as the most commonly used technologies in electric power systems and the uncertainties of demand forecasting are exploited in this work.

In [16], the Multi-Objective ORPD in a distribution network is presented for minimizing the active power loss, voltage deviation, and cost of reactive power compensators. In this paper, the effects of different load models on the cost functions are studied and a modified Grey-Wolf optimizer is used to solve the problem. In addition, the effects of installing DGs in the grid are investigated. However, the uncertainty of the DGs' power generation is ignored.

In [17], the effect of coordinated and automatic voltage regulation in a power system with wind turbines and SVC is presented. The objective function is mainly formulated as a single-objective reactive power optimization problem. In [18], the solution to the ORPD problem by the combined optimization method of particle swarm and various individual evolution is presented. The objective function of minimizing the active and reactive power losses of the network is considered. The obtained results were compared with the results of the PSO method.

The Multi-Objective SC-ORPD considering the technical uncertainties of power systems is presented in [19]. Maximizing the voltage stability margin based on the new voltage instability risk index and minimizing the losses are the objective functions. The SPEA-II optimization algorithm is then used to solve the problem. In [20], a novel risk-based ORPD algorithm with technical and economic objectives in a wind-integrated system is presented. The proposed method is more compatible with operational criteria. In this reference, a novel probabilistic-possibilistic risk index for risk quantification is presented.

In Table 1, the studied references are categorized according to their objective functions, risk-based or security-based planning, and their problem-solving algorithm. In Table 1, 68% of the studied works considered the problem as a multi-objective problem. 26% of the references considered contingencies in their problem (security constraint) and 47% of the references solved the problem with a risk-based formulation. It is noteworthy that, the contingencies are the main factors of the system risk. Thus, the subscription of the third and fourth columns shows the references that solved the problem using a risk-based formulation, and the risk is estimated after simulating the contingencies. Only 20% of these references belong to this category. 15% of the references consider all technical contingencies, including events, renewables, and load forecast contingencies. Finally, 52% of the references use a meta-heuristic algorithm to solve the problem.

C. The Necessity of Research

Based on the literature review and the results of Table 1, several existing gaps were identified, including:

1. A comprehensive objective function that risk and operational cost can be managed by ORPD or OPF simultaneously is not considered in the references.
2. The simultaneous consideration and management of the impact of risks arising from unplanned events or uncertainties on the power system and customers.
3. All technical uncertainties including RESs, load forecast, and contingencies are not studied in reviewed papers.
4. A contingency filtering method based on the riskiest events needs to be proposed. The proposed algorithm should be able to eliminate the influence of the control variables changing and updating during the optimization process.
5. PSO is a well-known, efficient and fast method in many engineering optimizations and planning. However, its premature convergence, lack of diversity, and being trapped in locally optimal solutions are mentioned as the shortage of this algorithm. Other optimization algorithms like NSGA and SPEA are accurate but with slow convergence.
6. *Novelty and main contribution*

   To address these issues, in this paper, a novel risk-based ORPD method is presented. The contributions of the are:

TABLE 1
COMPARING REFERENCES

Reference number	Multi-Objective	Security Constraint	Risk Based	All technical uncertainties	Meta-Heuristic
2	✓	×	×	×	×
3	×	×	✓	×	×
4	×	✓	✓	×	×
5	✓	×	✓	×	✓
6	✓	×	✓	×	×
7	✓	×	✓	×	×
8	×	×	×	×	×
9	✓	×	×	×	✓
10	✓	×	×	×	×
11	✓	✓	✓	✓	✓
12	×	×	×	×	✓
13	✓	×	×	×	✓
14	✓	×	✓	×	×
15	✓	✓	×	×	✓
16	✓	×	×	×	✓
17	×	×	×	×	×
18	×	×	×	×	✓
19	✓	✓	✓	✓	✓
20	✓	✓	✓	✓	✓

1- To cover 1 and 2, a comprehensive technical-economic objective function is proposed. The technical objective consists of voltage instability risk, voltage violation risk, and current flow violation risk, and the economic objective includes generator reactive power generation/absorption cost, active power loss, cost, load shedding cost, and generation compensation cost. The proposed method is a hybrid preventive-corrective approach.

The proposed method finds the optimal decision variables of generators, compensators, and transformers in such a way that there is no violation after the occurrence of uncertainties in the system. If this is not possible or costly, remedial actions (manual by the operators or automatic by the special protection systems) are performed after the occurrence of violations following a fault. In this way, the risks to the system (voltage instability, voltage violation, and power flow violation risks) and the risks to the customer (loss of load maintenance) can be simultaneously minimized or managed by assigning different important factors to the sub-objective functions by the operators or planners, depending on the network condition, system reactive power reserve.

Moreover, the operating limits of conventional and wind generators are considered in the problem formulation to increase the practicality of the proposed method.

2- To solve Problem 3, the uncertainties imposed on the system by the integration of renewable energy (doubly fed induction wind generators (DFIG) in this work), the uncertainty of load forecasting, and the probability of accidental equipment failure are included in the problem formulation of the proposed method by using a probabilistic approach to estimate the risk. For this purpose, the moment matching method (MMC) is used to generate optimal scenarios considering uncertain variable statistics. To reduce the size and select the best scenarios, the Fast Forward Selection (FFS) algorithm is used.

3- A new contingency filtering method to cover issue 4, is proposed that finds high risky contingency while considering the effect of the decision variable changing during the optimization procedure. In fact, the proposed method is integrated into the optimization process which is different from the usual methods that filtering is done before the optimization procedure is run.

4- To address 5, Hybrid Multi-Objective PSO with sine cosine acceleration coefficients (H- MOPSO-SCAC) is introduced and utilized based on H- PSO-SCAC to achieve Pareto-front of the solutions.

D. Organization of the Paper

The paper is organized as follows: Section II is devoted to the methodology. Subsection A describes how to model the random outage of the branches, wind and demand uncertainties, and scenario generation and reduction method using MMC and FFS, respectively. Subsection B presents, the proposed multi-objective risk-based ORPD. In subsection B, objective functions and constraints of the proposed method are described. Subsection C presents the proposed contingency filtering method and subsection D is devoted to a novel H-MOPSO-SCAC description. Section III is dedicated to the test results. First, the detail of the system under study is presented in subsection A. Simulation results are available in subsection B and subsection C is devoted to sensitivity analysis. Section IV is devoted to the conclusion.

II. Methodology

III. Power System Uncertainty Modelling

The operation and planning of power systems involve uncertainties. Loads, renewable resources, and random outages are the main sources of uncertainties in the grid. The probabilistic security assessment is a leading approach for dealing with these uncertainties and calculating the risk indices they cause. In this approach, the continuous probability distribution functions (PDF) of random variables are discretized using a limited number of scenarios. Then, the deterministic analysis is performed separately for each generated scenario and the expected values of the outputs are calculated.

1. *Random Branch Outage Probability*

The two-stage model, which represents a well-known approach to calculating the random probability of failure of the component, is used in this work [21]. In this model, the equipment can be in one of the states "in service" (available) or "out of service" (unavailable). The average unavailability of the equipment in a long-term time frame is calculated using [21]:

$$FOR = \frac{fr}{rr+fr} \quad (1)$$

The failure rate (FOR) can be used to estimate the failure probability of the random component.

2. *Modelling the Uncertainty of Wind Turbine Energy*

The energy generated by wind turbines is highly dependent on the wind, which fluctuates [12, 13]. It has been shown that deterministic modelling of fluctuating wind generators can lead to system instability. The most common PDFs for modelling wind speed uncertainty are Weibull or Rayleigh PDFs [23]. In this work, the Weibull PDF is used:

$$f(v)=\frac{k}{c}\left(\frac{v}{c}\right)^{k-1} exp\left(-\frac{v}{c}\right)^{k} \quad (2)$$

In this work, a linearized characteristic curve is also used to represent the relationship between turbine output power and wind speed. The data of the characteristic curve are available in [11].

3. *Demand Uncertainty Modelling*

Demand uncertainty due to load forecasting errors is usually estimated using the normal probability distribution function (PDF) [2]. In this paper, the mean of the normal PDF is assumed to be the forecast load, and the standard deviation is set to 3% [19].

4. *Scenario Generation and Reduction*

It is challenging to solve stochastic models that contain variables with continuous distribution. In such cases, a discretization process (scenario generation) is used to estimate the continuous PDF of the uncertain variable with a limited number of scenarios. However, in most cases, it is difficult or impossible to find the exact probability density function of random variables. Nevertheless, it is possible to approximate the continuous distributions using their statistical properties. Therefore, in this paper, the moment matching method (MMC) was used to create a limited number of scenarios. This method uses statistical properties instead of exact PDFs of the uncertain variables [22]. For this purpose, an optimization problem is defined that minimizes the distance between the statistical properties of the continuous and matched discrete distributions. The optimization problem is subject to the constraint that the sum of the relative importance of the statistical specifications (mean, variance, kurtosis, and skewness in this work) equals one. The MMC method is described in Appendix A.

However, increasing the number of uncertain variables leads to an increase in the number of scenarios and thus in simulation time. In this case, a strategy is needed to reduce the number of scenarios. This is done by selecting a subset of scenarios that represents the entire scenario (scenario reduction). Fast Forward Selection (FFS), a well-known algorithm for scenario reduction, is used in this work. The FFS procedures can be found in [22]. MMC is used to create 5 scenarios for wind speed and load forecasting (total number of scenarios= $[5]_{1*5}$* $[5]_{5*1}$). Then 6 scenarios are selected as reduced scenarios using FFS. Table 2 shows the reduced scenarios and their corresponding probabilities.

B. Proposed Multi-Objective Risk-Based ORPD

The general formula of the proposed risk-based ORPD with multiple objectives is presented in (3), which contains techno-economic objective functions (F1 and F2, respectively) that must be minimized under equality and inequality constraints (A, B).

$$\min \begin{cases} F1(vs,vc) \\ F2(vs,vc) \end{cases}$$

s.t

$$A(vs, vc)=b,$$

$$B(vs, vc)>0, \quad (3)$$

where, F_1 and F_2 are formulated as:

$$F_1=w_{11}*F_{11}+w_{12}*F_{12}+w_{13}*F_{13} \quad (4)$$

$$F_2=w_{21}*F_{21}+w_{22}*F_{22}+w_{23}*F_{23}+w_{24}*F_{24} \quad (5)$$

where, vs is the vectors of state variables, vc is the vector of control variables, w_{ij} are importance factor of sub-objective functions that are factors are assigned proportionally to the relative importance of each sub-objective and the different operational strategies. F11, F12, F13, F21, F22, F23, F24 are sub-objectives of technical and economical objective functions, respectively.

The vs and vc can be expressed as follows:

$$vs = [P_{slack}, Q_{slack}, V_L, Q_g] \quad (6)$$

$$vc=[V_g, Q_c, T_p, P_w, Q_w] \quad (7)$$

where, P_{slack}, Q_{slack} are active and reactive power of slack generator, V_L is the voltage of PQ buses, Q_g is the reactive power of generators. V_g is the voltage of generators, Q_c is the reactive power of compensators, T_p is the tap ratio of transformers, P_w, Q_w are the active and reactive power of wind generator, respectively.

1. Voltage Instability Risk (F_{11})

In most references, the risk index is quantized by the production of the consequence and probability of the events or uncertainties [9, 11, 14, 29]. For practical purposes, a certain level of violation and risk is assumed to be acceptable by the operators (for example, voltage deviation). The L-index which is presented in [23] is used in this paper to estimate the Voltage Stability Margin (VSM) following a contingency. So, using the L-index, a voltage instability risk index is developed which combines the probability and consequence of the events if VSM decreases from a pre-specified amount.

The index will become zero when an acceptable stability margin is achieved and otherwise increases linearly with decreasing VSM as follows:

$$VIR = \begin{cases} \sum_{n_l=1}^{n_l} \sum_{s=1}^{s_n} \frac{(L_i-VSM)}{L_{max}} * FOR_j & if\ L_i > VSM \\ 0 & if\ VSM > L_i \end{cases} \quad (8)$$

TABLE 2
SCENARIO REDUCTION RESULTS

Wind Speed (m/s)	LOAD %	Probability
0.436	1	0/123
3.52	1/011	0/17
4.7	0/988	0/174
1.283	0/949	0/127
4.7	1	0/226
10.5	1/053	0/18

2. Voltage Violation Risk (F_{12})

Similar to the voltage instability risk estimation, the voltage violation risk is calculated as follows:

$$VVR = \begin{cases} \sum_{n_l=1}^{n_l} \sum_{s=1}^{s_n} \frac{(V_i - EV)^2}{V_{min}} * FOR_j & if\ V_i > EV \\ 0 & if\ EV > V_i \end{cases} \quad (9)$$

3. Flow violation Risk (F_{13})

The flow violation risk can be calculated:

$$FVR = \begin{cases} \sum_{n_{br}=1}^{n_{br}} \sum_{s=1}^{s_n} \frac{SL_i - ER_i}{SL_i^{pre}} * FOR_j & if\ S_i > ER \\ 0 & if\ ER > S_i \end{cases} \quad (10)$$

4. Reactive power cost (F_{21})

Reactive power cost is due to increasing winding losses caused by the generation or absorption of reactive power by the generators, as well as the opportunity cost. The quadratic cost function can estimate reactive power cost as follows [10]:

$$EC_q = \sum_{s=1}^{s_n} \sum_{i=1}^{N_g} a_{gi}^q Q_{gi}^2 + b_g^q Q_{gi} + c_{gi}^q \quad (11)$$

5. Active power loss cost (F_{22})

The cost of active power loss can be estimated using the following equation [24]:

$$ELC = P^* \sum_{s=1}^{s_n} \sum_{k=1}^{N_{T-Line}} g_k [(V_i^2 + V_j^2 - 2V_i V_j cos(\delta_i - \delta_j)] \quad (12)$$

6. Load Cut Price (F_{23})

When an emergency occurs, following an event usually, if certain operational limits are violated for a short period, some loads can be manually or automatically cut as a remedial action. Special Protection Systems (SPS) are capable devices to perform automatic load shed in such cases. These devices are designed for multiple stages. The cost of load cut can be expressed as follows:

$$ELSC = \sum_{s=1}^{s_n} \sum_{i=1}^{NC} \sum_{j=1}^{n_{stage}} \sum_{f=1}^{n_{lc-load}} P_j L_f * FOR_{NC} \quad (13)$$

7. Generation Compensation Cost (F_{24})

In power system operation, if there is a generator failure or a deviation from the expected power of the wind turbine, the remaining plants must compensate for the active power generation and consumption unbalance to maintain the frequency at its predefined value. In the planning stage, if the imbalance between power generation and consumption of active power exceeds the capability of the slack bus to generate power, the power flow may diverge. The power flow divergence may be considered an instability by mistake. To avoid such situations, the active power imbalance must be compensated in active power unbalance cases. In certain ISO policies, this compensation is paid to the plant with an additional fee. This can be expressed as follows:

$$EGCC = \sum_{s=1}^{s_n} \sum_{i=1}^{NC} \sum_{j=1}^{ng} PF_j * \Delta P * IC_j * FOR_{NC} \quad (14)$$

8. Power Flow Constraints

Power flow equations are equality constraints of the OPF problem.

$$P_{G_i} - P_{D_i} - V_i \sum_{j=1}^{Nb} V_j [G_{ij} \cos(\delta_i - \delta_j) + B_{ij} \sin(\delta_i - \delta_j)] = 0 \quad (15)$$

$$Q_{G_i} - Q_{D_i} - V_i \sum_{j=1}^{Nb} V_j [G_{ij} \sin(\delta_i - \delta_j) + B_{ij} \cos(\delta_i - \delta_j)] = 0 \quad (16)$$

9. Voltage Constraints

The voltage of the controlled and uncontrolled buses, including wind turbine buses, must be within the highest and lowest limits in the base case.

$$V_i^{min} \le V_i \le V_i^{max} \quad (17)$$

10. Reactive Power Constraints

If the reactive power output of the generators exceeds the capability curve range, the bus voltage will become uncontrollable and the bus will be converted from PV to PQ in power flow studies. The operating point, internal voltage, armature current, and reactance of the generators are factors that influence the capability curve of the generators. Using these factors, the capability curve of the generators and wind turbines can be estimated. The formulas for modelling generators and wind turbine capability curves can be found in [11]. They are not provided here to keep the text concise.

11. Transformer Constraints

The higher and lower bounds of tap number of transformers are as follows:

$$T_i^{min} \le T_i \le T_i^{max} \quad (18)$$

12. Shunt Compensators limits

The VAR generation/absorption of the shunt compensator is limited by its lower and upper bounds.

$$Q_{ci}^{min} \le Q_{ci} \le Q_{ci}^{max} \quad (19)$$

13. Security Constraints

Transmission line loadings and voltage stability margin limits must be met in the base case. They can be defined as follows:

$$S_{li} \le S_{li}^{max} \quad (20)$$

$$VSM_i < VSM_{pre-def} \quad (21)$$

C. Contingency Screening

The contingency filtering algorithm separates critical events from non-critical ones to be simulated, thereby reducing simulation time. The contingencies from the events list that lead to high risk are considered critical. In the proposed method, the advantage of population-based optimization algorithms will be utilized to perform a sensitivity analysis and filter high-risk contingencies. Indeed, the screening algorithm is integrated into the optimization problem for this purpose and will be updated during the iterations. The advantage of the proposed approach is that it allows for the examination and tracking of the effect of variations in the control variable on the system's security, reliability, and customer satisfaction. The algorithm is as follows:

Step 1: Create the initial contingency list.

Step 2: Create an initial random population.

Step 3: Using the initial contingency list, perform contingency analysis for all population members and identify the events with the highest risk.

Step 4: Eliminate any repetitive events from the list and create a list of critical events.

Step 5: Update the population and continue the optimization procedure with the selected list of critical events.

Step 6: If the optimization iteration reaches the recheck number, replace the critical events list with the initial contingency list.

Step 7: Go back to step 3 and continue until the optimization stop criteria are met.

The list containing high-risky events is replaced by the initial contingency list in certain optimization iterations to check and identify the most severe events with variations in control variables.

D. H-MOPSO-SCAC Optimization Algorithm

PSO is a well-known and efficient method for many engineering optimization and planning problems. The advantages of the PSO method are that it is easy, simple, and can be implemented in any working environment. However, premature convergence, lack of diversity, and being trapped in locally optimal solutions are mentioned as shortcomings of the PSO [25]. To address these limitations, a novel search strategy is proposed that combines the hybrid particle swarm optimizer (H-PSO) with sine cosine acceleration coefficients (SCAC) within the PSO framework and is called H-PSO-SCAC [26]. Based on modifications [26], this paper introduces a modified version of multi-objective PSO (MOPSO) called H-MOPSO-SCAC, which is designed to handle multi-objective problems. These modifications are as follows:

1- Opposition-based learning (OBL) can increase the opportunities of reaching the global optimal solution.

i- Randomize the initial population.

ii- Calculate the reverse population based on (22).

$$X_{max.j} + X_{min.j} - X_{ij} \quad (22)$$

Where, $X_{max.j}$ and $X_{min.j}$ are the population position max and min value at dimension j, respectively, X_{ij} is the population member i and $X'_{ij}$the opposite population.

v- Using domination theory, select the nondominated solution between X_{ij} and X'_{ij} to maintain the population size.

2- The sine map is used to adjust the inertia weights ω of the PSO. The range of the sine map is [0,1]. The inertia weight ω is given as follows:

$$\omega = \frac{c}{4}\sin\left(\pi_{x_{k-1}}\right) \quad (23)$$

Where, k is the current iteration number.

3- A modified update formula for positions in each iteration is used. This modification can improve the exploration and exploitation of the algorithm.

$$X_{i+1}^{d} = X_i^d * W_{ij} + V_i^d * W'_{ij} + \rho * gbest^d * \varphi \quad (24)$$

Where, W_{ij} and W'_{ij} are the dynamic weights, $gbest$ is the current global best, X_i^d and V_i^d are the previous solution and velocity respectively, ρ random number between 0 and 1, φ is the coefficient that determines the maximum step size.

$$w_{ij} = \varphi = \frac{\exp\left(\frac{f(j)}{u}\right)}{1+\exp\left(-\frac{f(j)}{u}\right)^{iter}} \quad (25)$$

$$w'_{ij} = 1 - w_{ij} \quad (26)$$

Where, u is the mean fitness value in the first iteration, iter is the current iteration and f(j) is the fitness of the jth particle.

4- The sine cosine acceleration coefficients are:

$$c_1 = \sigma * \sin\left(\left(1 - \frac{M_j}{M_{max}}\right) * \frac{\pi}{2}\right) + \delta \quad (27)$$

$$c_2 = \sigma * \cos\left(\left(1 - \frac{M_j}{M_{max}}\right) * \frac{\pi}{2}\right) + \delta \quad (28)$$

Where, σ and δ are the constant (σ=2, δ=0.5)

Fig.1 shows the flowchart of the proposed method and the optimization procedure.

III. TEST RESULTS

A. System Under Study

The proposed method is applied to the modified IEEE 30-bus test system. The test system is modified by installing a 60 MW Doubly Fed Induction Generator (DFIG) wind farm at bus 20. In addition, the buses and branches of the network are equipped with 2-stage under-voltage and flow violation SPS systems. In each stage, the SPS system will reduce the load on the bus by 10%. The minimum VSM is assumed to be 40%, high and low voltage $\pm 5\%$ and maximum loading of the branches is considered 90%. Fig. 2 represents the modified

network. The simulations are conducted using Manpower

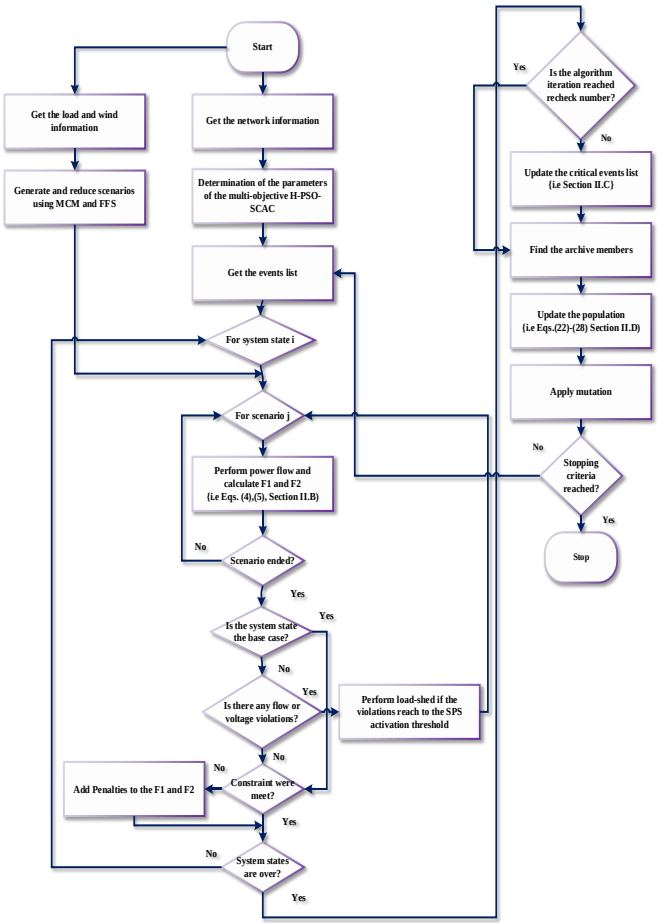


Fig. 1. The flowchart of the proposed method

version 6 software packages. The algorithm was implemented in MATLAB R2017a and executed on a core-i5 laptop with a 2.6 GHz processor and 6.0 GB of RAM. Necessary data for calculating the costs of the generator's active and reactive power, as well as the system data, are available in Appendix B. The PSO population size algorithm is set to 60 and the repository size is set to 25. Maximum iteration is set to 100.

B. Simulation Results

To evaluate the effectiveness of the presented method, three scenarios are investigated as follows:

1- The ORPD considering wind and load uncertainties, expected L-index and EL as F1 and F2.
2- SC-ORPD
3- Proposed RB -ORPD *scenarios*
4- Sensitivity analysis of the method

1. Scenario I: ORPD considering wind and load uncertainties, expected L-index and EL as F1 and F2.

This scenario is simulated to compare the results of the proposed method with those of other literature. Table 3 displays the Pareto front for the proposed method, other references, and the original MOPSO approach. F1 for the

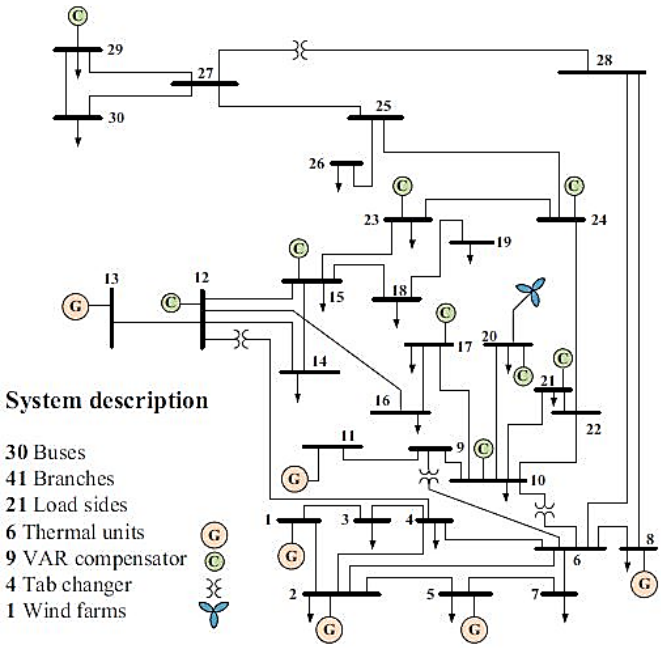


Fig. 2. The single-line diagram of the test system

2. *Scenario II: SC-ORPD*

proposed method ranges from 0.0998 to 0.151, and F2 ranges from 3.23 to 4.09 in the Pareto front. It is evident from the results that the proposed method outperforms other methods in objective minimization, except for reference [11] in the minimization of F1. The results obtained by reference [11] are 6% better in terms of F1 and 13% worse in terms of F2 compared to the proposed method in the objective minimization case. However, the simulation time of the proposed method is 28% of that in [11]. Comparing the obtained results to those of the original MOPSO demonstrates the superiority of the proposed method in terms of objective minimization and solution diversity. Although the simulation time of the proposed method is 56% longer than that of the original MOPSO due to additional calculations.

This scenario is conducted to analyze the impact of security rather than risk-based planning. The L-Index is considered F1, and the economic objective is considered F2. The voltage of the buses and the flow limits are formulated as hard limits. The proposed method for filtering events has been modified to select the most severe contingencies instead of high-risk ones. To investigate the effect of the SPS system, a simulation was conducted under heavy load conditions (load factor = 1.1). Also, two operational strategies with different weighting factors for the sub-objectives are investigated to analyze the effect of the relative importance of each function. The load cut price in Stages 1 and 2 are set at 10 and 20 $/MW, respectively. The incremental and opportunity cost of the generators is considered to be 10 $/MW. The initial contingency list contains the following branch numbers: 4, 39, 38, 36, 13, 19, and 14, as well as generator numbers 3, 4, 5, 6, and wind farm outages.

TABLE 3
COMPARING THE RESULTS WITH PUBLISHED REFERENCES

Reference	EL-INDEX	EL (MW)	TIME
Reference [27]	0.1192-0.1317	4.55-7.07	-
Reference [28]	0.111-0.121	4.4-6.6	-
Reference [29]	0.1161-0.1258	3.4815-4.0646	-
Reference [11]	0.0939 -0.17	3.74-4.082	4000
Original MOPSO	0.1101 -0.1208	4.35-4.73	722
Proposed method	0.0998-0.151	3.23-4.09	1127

The results are available in Table 4. By employing the first strategy (using similar weighting factors for the sub-objective functions), F1 varies from 0.79 to 0.799, and F2 varies from 138.2 to 196.63. All solutions involve 4.6 MW of load shedding. In the second strategy, the F1 varies from 0.77 to 0.808 and the F2 varies from 95.5 to 165 dollars.

In this strategy, the load cut of the solutions includes 0, 4.6, and 10.9 MW. The solution with the lowest cost and highest security is related to 0 MW of load shedding. With this selection, the customer's power supply is maintained. In the first strategy, the scheduled VSM is preserved by implementing load cuts for all solutions. However, in the second strategy, some solutions result in minor VSM violations.

By assigning different weighting factors, operators can mitigate the effects of load increases and events on the system and customers. As evident from the results, the operational cost has significantly increased due to the network's heavy load condition and the network's security-focused planning.

3. Proposed RB-ORPD

In this scenario, the proposed method based on the flowchart shown in Figure 3 is applied to the test system. The likelihood and consequences of the event are used to calculate the risk of contingencies. The scenario is divided into cases with two weighting factor strategies, similar to the previous scenario. Table 5 shows the results. As is clear from the results, RB-ORPD provides fewer economic objectives compared to Scenario 2. The branches have a lighter load and the bus voltages are at a higher level. Hence, not only does the economic objective decrease, but the security of the system also increases compared to traditional planning.

In strategy 1, load shedding of 0, 4.6, and 12.7 MW is performed in the solutions. While strategy 2 leads to 0, 4.6 MW of the load is shed. In Strategy 2, the emphasis is on utilizing the reactive power of the generators and shunt compensators. So, a reduction in load shedding is implemented, and the economic objective is decreased. So, it can be inferred that risk-based planning, compared to traditional planning, will lead to better and more efficient allocation of resources. The economic objective is achieved and the security of the system is also increased. In addition, the impact of these events on the customers is also reduced.

TABLE 4
THE RESULTS FOR THE SCENARIO II

WEIGHTING FACTORS W21-W22-W23-W24	F1	F2 ($)	MAXIMUM VSI	MAXIMUM LOADING %	MINIMUM VOLTAGE P.U
1-1-1-1	0.79-0.799	138.4-141	0.79	111	0.94
0.2-0.2-0.3-0.3	0.77-0.808	95.5-165	0.77	118	0.91

TABLE 5
THE RESULTS FOR THE SCENARIO III

WEIGHTING FACTORS W21-W22-W23-W24	F1	F2 ($)	MAXIMUM VSI	MAXIMUM LOADING %	MINIMUM VOLTAGE P.U
1-1-1-1	0.0145-0.155	93-153.5	0.7922	1.02	0.973
0.2-0.2-0.3-0.3	0.0148-0.1625	48.34-101	0.7937	1.0234	0.954

4. Sensitivity Analysis

This subsection is dedicated to the sensitivity analysis of the proposed method. First, the sensitivity analysis is done on the MMC method and next the effect of assigning the various important factor to the sub-objective functions is investigated.

4.1 *MMC method sensitivity analysis*

Since using different importance factors of statistical properties leads to various optimal solutions, in this sub-section different importance factors are assigned to examine the sensitivity of the method to these parameters and the robustness of the method. In addition, the sensitivity to the number of statistical properties is also analyzed. Table 6, shows the obtained results. It can be concluded from the obtained results that, the effect of various important factors on the OFV depends on the opinion of decision-makers and planners. The OFV reaches near zero in all cases. However, in the cases where a statistical property is assigned a greater importance factor, the number of estimated outcomes by the algorithm may decrease. In fact, the same scenarios may be estimated. In the case with $\omega_1 = 1$, 10 identical points equal to the mean value will be estimated by the MMC. In addition, if the number of outputs decreases, the OFV will be increased. For the case with 2 outputs, the OFV reaches 0.142 which is unacceptable. Similar results will be achieved if the PDF is replaced with other PDFs like the normal probability distribution function.

4.2 *Various important factors assigned to sub-objective function F2*

Table 7 represents the obtained results for this sub-section. Various important factors are assigned to the sub-objective function of the F2 to investigate their effects on the obtained

results. The simulation is done in the heavy load condition. First, the important factor of the F_{21} is set to zero, and the others 0.33. Therefore, the cost of reactive power provided by the generators is assumed to be free. The results show that compared to the cases in this case, the technical and economic objective is placed between other cases. The solution of the Pareto front has load cuts of 0, 4.6 MW.

In the second case, the important factor of the F_{23} is set to zero and the others are set to 0.33. In this case, there is no charge for the load cut. So, all the solutions of the Pareto front have load shed from 4.6 to 22.27 MW. Compared to other cases in this case minimum F1 and F2 are achieved. Hence, it can be expressed that the F1 (risk functions) are more sensitive to the load in the simulation condition of the network compared to the reactive power generation of the generators.

This means that, in the PV analysis of the network, the operating point of the network may be close to the nose point. Security analysis of the network shows that without performing reactive power optimization, the system may experience instability following contingencies in the heavy load condition. The system side risk is decreased and the customer side is increased instead in this case.

In the third case, the F_{24} is set to zero and the others to 0.33. It means that the system may have the highest active power balance compared to the other cases. The results show that compared to case 1, F1 and F2 have decreased. In addition, compared to case 1, the solution with a zero MW load cut is achievable. Therefore, it can be concluded in the simulated network operating condition, the active power providing may increase system security compared to reactive power.

IV. Conclusions

In this paper, a multi-objective and risk-based optimal reactive power dispatch with a techno-economic objective function is proposed. The problem formulation takes into account the uncertainty of the wind turbine energy and load forecast. The practical limits of conventional and wind generators were modelled to estimate the capability curve of the generators. In addition, special protection systems were modelled in the proposed method. Based on Hybrid-PSO with sine cosine acceleration coefficients (H-PSO-SCAC), the original multi-objective PSO (MOPSO) was modified, and H-MOPSO-SCAC was presented. The obtained results demonstrated that:

1- The proposed H-MOPSO-SCAC provides better solutions in objective minimization and solution diversity compared to the original PSO and previously published methods. In addition, the simulation time of the proposed method is reasonably low compared to the other methods except the original MOPSO.

2- The analysis of the results showed that considering contingencies in the planning stage (SC-ORPF), may increase operational costs. Also, ignoring the contingencies may lead to system instability (as it is clear from comparing Table 3 and Table 4 results that VSM may decrease after the events significantly).

TABLE 6
THE SENSITIVITY ANALYSIS OF THE MMC METHOD

	IF	N.Out	OFV	IF	N.Out	OFV
ω_1	0.25	10	$1.0 * 10^{-18}$	0.55	10	$1.4 * 10^{-18}$
ω_2	0.25			0.15		
ω_3	0.25			0.15		
ω_4	0.25			0.15		
	IF	N.OUT	OFV	IF	N.OUT	OFV
ω_1	0.15	10	$2.2 * 10^{-21}$	0.15	10	$5.3 * 10^{-17}$
ω_2	0.55			0.15		
ω_3	0.15			0.55		
ω_4	0.15			0.15		
	IF	N.OUT	OFV	IF	N.OUT	OFV
ω_1	0.15	10	$5.7 * 10^{-17}$	0.5	10	$1.6 * 10^{-22}$
ω_2	0.15			0.5		
ω_3	0.15			0		
ω_4	0.55			0		
	IF	N.OUT	OFV	IF	N.OUT	OFV
ω_1	0.5	10	$2.9 * 10^{-17}$	1	10	$1.8 * 10^{-18}$
ω_2	0			0		
ω_3	0.5			0		
ω_4	0			0		
	IF	N.OUT	OFV	IF	N.OUT	OFV
ω_1	0.25	100	$1.1 * 10^{-17}$	0.25	2	0.142
ω_2	0.25			0.25		
ω_3	0.25			0.25		
ω_4	0.25			0.25		

TABLE 7
VARIOUS IMPORTANT FACTORS FOR THE F2

CASE NO	WEIGHTING FACTORS W_{21}-W_{22}-W_{23}-W_{24}	F1	F2 ($)	LOAD CUT (MW)
1	0-0.33-0.33-0.33	0.0146-0.162	43-56	0-4.6
2	0.33-0.33-0-0.33	0.0137-0.152	27.6-31.6	4.6-10.97-22.27
3	0.33-0.33-0.33-0	0.0144-0.153	35-50	0-4.6-10.97

3- Risk-based ORPD led to fewer economic objectives and more security compared to SC-ORPD planning. Using both planning methods, the desired voltage stability margin was achieved. However, less load cut was performed and bus voltages were increased and branch flows were decreased in risk-based planning.

4- In addition, the results revealed that the solutions with the lowest load shed (lowest customer side risk) will lead to the highest system side risk and vice versa. However, the solution that compromises between the customer and system side risk is achievable using the proposed formulation and assigning different weighting factors to the objective functions. The operators can to select the solution with a high loss of load expectation (high customer risk) and lower system risk or vice versa according to the future operation condition or

planning strategies.

The author's proposal for future research in this field are:

1- A better and more efficient risk function must be introduced to estimate the real risk of the events. Usually, FOR of the network components in the power system is near zero. Hence, the risk associated with power system component failure may be estimated as insignificant. Especially, in high-impact and low-probability event cases.

2- Probable modelling of protection relays can increase the applicability and practicality of the planning.

3- Considering the risk of not realizing the prediction made about the occurrence of random variables in the planning.

REFERENCES

[1] J.D. McCalley, V. Vittal, N. Abi-Samra, "An overview of risk based security assessment," IEEE Power Engineering Society Summer Meeting. Conference Proceedings Edmonton, Alta., Canada, 1999.

[2] A. Rabiee, A. Soroudi, B. M. Ivatloo, "Corrective voltage control scheme considering demand response and stochastic Wind power," IEEE Transactions on Power Systems, 2014.

[3] A. Rabiee, A. Souroudi, A. Keane, "Risk-Averse preventive voltage control of AC/DC power systems including wind power generation," IEEE Transaction on Sustainable Energy, 10.1109/ TSTE. 2451511, 2015.

[4] L. Roald, M. Vrakopoulou, F. Oldewurtel, G. Andersson, " Risk-Constrained optimal power flow with probabilistic Guarantees," Power Systems Computation Conference, 2015.

[5] Y. . Li, Q. H. Wu, "Downside riskconstrained probabilistic optimal power flow with wind power integrated," IEEE Transactions on Power Systems, Volume: 31, Issue: 2,2016.

[6] L. Roald , G. Andersson, S. Misra, M. Chertkov, S. Backhaus, "Optimal power flow with wind power control and limited expected risk of overloads," Power Systems Computation Conference (PSCC), Genoa, Italy,2016.

[7] L. Roald, S. Misra, A. Morrison, G. Andersson, "Optimized risk limits for stochastic optimal power flow," IEEE 56th Annual Conference on Decision and Control (CDC) December 12-15, Melbourne, Aus,2016.

[8] M. Nojvan, H. Seyedi, B. H. Ivatloo, "Preventive voltage control scheme considering demand response, correlated wind and load uncertainties," Journal of Energy Management and Technology (JEMT), 13, Paper no. JEMT1705-1011, 2017.

[9] P. Fakhrooeian, M. Abedi, P. Karimyan, "optimal allocation and sizing of dynamic VAR support to improve short-term voltage stability considering wind farm and dynamic load model," International Journal of Industrial Electronics, Control and Optimization, Vol. 1, No. 1,pp. 41-51, June 2018.

[10] M. Ghaljehei, Z. Soltani, J. Lin, G.B. Gharehpetian, M.A Golkar, "Stochastic multi-objective optimal energy and reactive power dispatch considering cost, loading margin and coordinated reactive power reserve management," Electric Power Systems Research 166, 163–177, 2019.

[11] M. Mokari, M.H. Moradi, "Security constraint optimal reactive power dispatch under uncertainty in a wind integrated power system," 55th International Universities Power Engineering Conference (UPEC),2020.

[12] M. Knittel, N. Majumdar, M. Schneider, N. Thie, A. Moser, "Voltage control in transmission grids considering uncertainties of renewable energy sources," 6th IEEE International Energy Conference, 2020.

[13] M. Ettappan, V. Vimala, S. Ramesh, V. Thiruppathy Kesavan, "Optimal reactive power dispatch for real power loss minimization and voltage stability enhancment using artificial Bee colony algorithm," Microprocessors and Microsystems doi.org/10.1016/j.micpro.2020.103085, 2020.

[14] P. Lei, C. Wang, Q. Wu, M. Yan, "Risk-based distributionally robust real-time dispatch considering voltage security," IEEE Transaction on Sustainable Energy, Volume: 12, Issue: 1 , 2021.

[15] M. Hassan, S. Kamel, M. El-Dabah, T. Khurshaid, J. L. Dominguez-Garcia, "Optimal reactive power dispatch with time-varying demand and renewable energy uncertainty using Rao-3 algorithm," IEEE Access (Volume: 9), DOI: 10.1109/ACCESS.2021.3056423.2021.

[16] S. Hosseini-Hemati, S. Karimi, G. H. Sheisi, "Multi-Objective ORPD considering different load models for active distribution networks," International Journal of Industrial Electronics, Control and Optimization, Vol. 4, No. 2,pp. 191-210, April 2022.

[17] V. Zakutsky, M. G. Gadzhiev, "Coordinated voltage and reactive power control of power systems with wind farms and SVCs as problem," 4th International Youth Conference on Radio Electronics, Electrical and Power Engineering, 2022.

[18] M. S.Saddique, S. Habib, S. S. Haroon., A. R. Bahatti, S. Amin, "Optimal solution of reactive power dispatch in transmission system to minimize power losses using sine-cosine algorithm," IEEE Access, Vol: 10, 2022.

[19] M.H. Moradi, M.Mokari, M. Abedini, "An investigation of technical uncertainties in power networks with the aim of improving voltage stability and reducing system losses using the SPEA-II optimization algorithm," Iranian Electric Industry Journal Quality & Productivity Vol. 11 / No. 1/ Spring 2022 (In Persian).

[20] M.Mokari, M.H.Moradi, M. Abedini, "Fuzzy risk-based optimal reactive power dispatch in a wind-integrated power system," Tabriz Journal of Electrical Engineering (TJEE), vol. 53, no. 1, Spring 2023.

[21] W. Li, "Risk assessment of power systems: Models, methods, and applications, Second Edition," John Wiley & Sons, 2014.

[22] Z. Hu, GA. Hu, "Two-stage stochastic programming model for lot-sizing and scheduling under uncertainty," International Journal of Production Economics, 180 , 198{207}, 2016.

[23] P. Kessel, H. Glavitsch, "Estimating the voltage stability of a power system," IEEE Transactions on Power Delivery, 346-54, 1986.

[24] M. Jazaeri, M. Mokari, "A new technical-economical algorithm to improve asymmetry in system voltage by SVC," European Journal of Scientific Research. vol. 115 No.3. 2013.

[25] C. A. Coello, G. T. Pulido, M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," IEEE Transactions on Evolutionary Computation, VOL. 8. 8, NO. 3, 2004.

[26] K. Chen, F. Zhou, L. Yin, Sh. Wang, Y. Wang, F. Wan, "A hybrid particle swarm optimizer with sine cosine acceleration coefficients," Information Sciences 422, 218–241, 2018.

[27] A. Abou El Elaa, M.A. Abidob, S.R. Spea, "Differential evolution algorithm for optimal reactive power dispatch," Electric Power Systems Research 81 (2011) 458–464.

[28] A. Ghasemi, Kh. Valipour, A. Tohidi, "Multi objective optimal reactive power dispatch using a new multi objective strategy," Electrical Power and Energy Systems 57 (2014) 318‑334.

[29] S. M. Bonab, A. Rabiee, B. M. Ivatloo, "Multi-Objective optimal reactive power dispatch considering uncertainties in the wind integrated power system," Springer international Publishing , 10.1007/978-3-319-51118-4_12,AG 2017.

Appendix A

Moment Matching Method

In this paper, the Moment Matching (MMC) method is employed for discretizing continuous PDF of uncertain variables. In this method, an optimization problem must be solved that minimizes the distance between statistical specifications of continuous distributions and statistical properties of fitted discrete distributions subject to a constraint ensuring the summation of the weighting factors to be one. In this study mean, variance, kurtosis, and skewness have been used as statistical specifics.

$\min \sum_{k\epsilon K} \omega_k (f_k(x,\pi)\text{-}VAL_k)^2$

s.t.

$\sum_{v\in\gamma} \pi_v = 1$

$\pi_v \leq 0 \; v \in \gamma$

Where, k is a statistical property, K is a set of all statistical properties, ω_k refers to the importance weight of statistical property k, v is the outputs, x is the vector of realizations for uncertain factors, π Probability vector for outputs, VAL refers to the value of the statistical property of k. We assume the continuous distribution of uncertain factors are independent of each other. The minimum number of outcomes can be achieved by:

$(D+1)y\text{-}1 \sim$ the number of scenarios

Where, D indicates the dimension of the scenario, and y represents the number of outputs. In this paper D is 2 (the product of 2 random variables and 1 period of time). The number of specifications is 8 (the product of 2 uncertain variables, 1 period of time, and 4 moments). Therefore, according to (8), the value of y is 4. To get better results, we chose 5 as the number of scenarios for each random variable.

The nonconvex nonlinear problem was solved by the COUENNE solver in GAMS 23.5. The weighting factors are considered to be equal. The obtained solution with an objective function value of zero or close to zero demonstrates that generated outcomes have statistical properties which perfectly match with specified properties of continuous distributions.

Appendix B

The proposed method is implemented on the modified IEEE-30 bus system. This network has 30 buses, 6 generators (bus 1 as slack bus and 2, 5, 8, 11 and 13 as PV buses), 41 branches (including 4 ULTC transformers at buses 6-9, 6-10, 4-12, 28-27 and non-transformer branches) and 9 compensators (buses 10, 12, 15, 17, 20, 21, 23, 24 and 29). A 60 MW DFIG wind farm is installed at bus 20. The minimum and maximum voltage control range of the ULTCs and the generators are set to 1.1-0.9 P.U. Table A, represents the active and reactive cost data used for the network.

TABLE A
PARAMETERS FOR ACTIVE AND REACTIVE POWER COST

Generator	a_p	b_p	c_p	a_q	b_q	c_q
G_1	0.02	2	0	0.0084	-0.00075	0.2
G_2	0.0175	1.75	0	0.007	0.00322	0.84
G_3	0.0625	1	0	0.0073	-0.00344	0.89
G_4	0.00834	3.25	0	0.0073	-0.00344	0.89
G_5	0.025	3	0	0.0073	-0.00344	0.89

Meysam Mokari was born in Hamedan, Iran. He received his B.Sc degree in Electrical Engineering from Islamic Azad University, Hamedan, Iran, in 2008, and his M.Sc. and Ph.D. degrees in Electrical Engineering from the Islamic Azad University of Saveh, Iran and Bu-Ali Sina, Hamedan, Iran, in 2011 and 2023, respectively. From 2013 he has been with Iran Grid Management Company (IGMC), Tehran, Iran. His current research interests include power system operation and control, power system planning, optimization and renewable energy systems.

Mohammad Hasan Moradi received the B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran; Tarbiat Modarres University, Tehran; and Strathclyde University, Glasgow, Scotland, in 1991, 1993, and 2002, respectively. He is a professor at Bu-Ali Sina University, Hamedan, Iran. His current research interests include new and green energy, microgrid modelling and control, power electronics, combined heat and power plant, power quality, supervisory control, and fuzzy control.

Mohammad Abedini was born in Boroujerd, Iran, in 1986. He received his B.Sc degree in electrical engineering from the Islamic Azad university of Borujerd, Borujerd, Iran in 2006. M.Sc. and Ph.D degrees in electrical engineering from Bu-Ali Sina University in 2008 and 2010,respectively. He is currently an associate professor at ayatollah Borujerdi university. His current research interest includes microgrids and distribution systems.