Virtual Sample based Multi-objective Optimization for High Pressure Turbine Disc

Guoqi Feng, Haiying Wang
School of Business Administration, Northeastern University, Shenyang, Liaoning
Email: Gqfeng@mail.neu.edu.cn; wanghy@mail.neu.edu.cn

Dongliang Cui
State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, Liaoning
Email: cuidongliang@mail.neu.edu.cn

Abstract. Mass and radial deformation are of great importance for high pressure turbine disc (HPTD), but computational cost of CAE (computer aided engineering) is too high to optimize these mutually restricted objectives. A virtual sample based method is proposed to speed the optimization of HPTD: noised based virtual samples are implemented to enlarge the training set, a cost-effective back propagation neural networks (BPNN) is trained whose hidden layer is set according to noise intensity and size of training set; this BPNN is used as fitness function of genetic algorithm for optimization whose initial population is the combination of different sample sets. Experiment shows that this data-driven framework decreases the engineering difficulty of MOO (multi-objective optimization), and it has high popularization value to optimization of other complex products.

1 .INTRODUCTION

In the development of modern aero engine, the most expensive component is high-pressure turbine because of the harsh working environment (Joly et al.2013). Mass of a HPTD maybe ten times of that of a low pressure turbine disc and this mass determines the efficiency, stability, service life and performance of the engine (Jahed et al. 2005), much work has been done on mass reduction of HPTD. The disc expands in presence of higher temperature and contracts in presence of lower temperature; the centrifugal force is yielded by a high rotational speed. The radial deformation of the disc is so influenced by these variables creating the problem of blade-tip clearance (Hailia et al. 1982): Tighter clearance reduces air leakage over blade tips, but over-tighter clearance causes rubbing between blades and shroud. Proper blade-tip clearance results in economic and environmental benefits to the public at large, most commercial aircraft engines take active thermal control for this tip clearance (Decastro et al. 2004). If deformation of disc is taken into account at design stage, difficulty of active clearance control will be reduced in flight, but it is difficult to design a disc with minimum mass and radial deformation because there is no optimal solution for both sides.

FEM is a widely used technique for modeling complex systems in computational engineering domain. However, analyzing of complex systems using FEM involves the application of time consuming methods of numerical simulation. Surrogate-assisted evolutionary computation is a good choice for this time-consuming problem (Jin 2011).Neural network has stronger ability of universal approximation than other surrogates, and many engineering use FEM-based NN as fitness function for MOO problems. Nagendra et al. (2005) developed a MDO system for turbine disc with the integration of NN and FEM. Li and Water et al. (2015) developed a design environment which helped engineers to automatically model, simulate and optimize transformer design using ANN and FEM. Bakhtiari and Karimi et al. (2016) used ANN to model twist extrusion process based on results of FEM, and optimization was carried out by a combination of NN and multi-objective meta-heuristic optimization algorithms. But computational load may still be a barrier to generate enough samples to train NN for complex product. With the inspiration of "virtual samples"

(Niyogi et al. 1998) and "The addition of noise to the input data of a neural networks during training can lead to significant improvement in generalization performance" (Bishop 1995), noise based virtual samples technologies could be used to enlarge sample size for surrogate model. However, traditional methods can't be directly applied on DOE samples because they are based on distribution information of incomplete data (Li et al.2015).

The rest of paper is organized as follows: Section 2 briefly describes the HPTD design details and main factors are selected for optimization; an artificial neural network (ANN) with high cost-performance is trained in Section 3; NSGA-2 algorithm with initial generation design are formulated in Section 4, And in section 5 certain turbine disc is optimized by the combination of ANN and NSGA-2. Finally, conclusions and future work will be summarized in Section 6.

2. SENSITIVITY ANALYSIS FOR HPTD DESIGN

As can be seen from Figure.1a, HPTD is a product with complex structure, and it is difficult to make a fully analysis because of its complicated working conditions. With considering the features of axial symmetry, 2D FEM could be used to reduce the cost of CAE.

2.1 CAE MODELING OF HPTD

Turbine disc consists of body, mortises and flanges. Body consists of rim, hub and web plate that are connected by different kinds of profiles; blades are fitted onto disc by mortises; flanges connect disc with other components. Information of mortises is fused into rim to make an equivalent "body-flange" model for 2D-FEM.

Figure.1b~c describe geometric details of certain HPTD, body parameters are rim radius R1, thickness W1 and height H1, shaft radius R2, hub thickness W2 and height H2, web plate's outer diameter R3 and inner diameter R4, outer thickness W3 and inner thickness W4; flange parameters are location (H3,W5,H5,W7), length (H4,H6-H5), width (W6,W8) and chamfer (R5,R6).

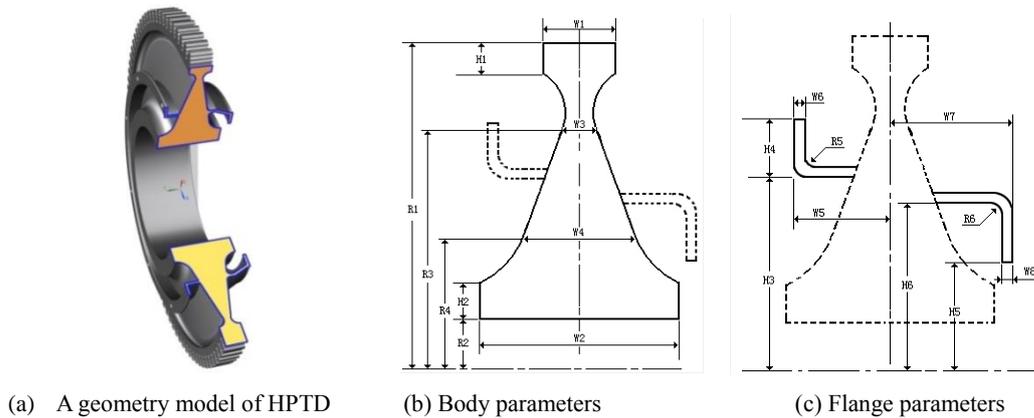| (a) A geometry model of HPTD | (b) Body parameters | (c) Flange parameters |

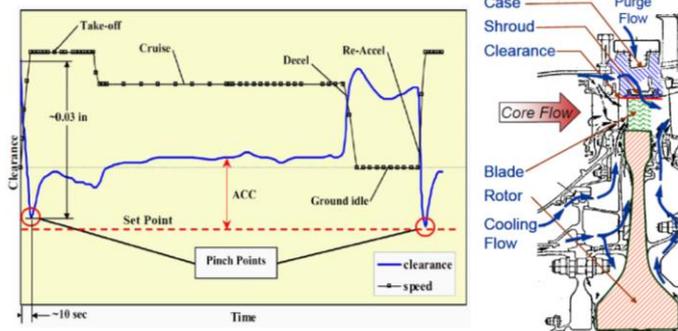Figure 1: geometry model based on "body-flange" structure

Figure 2: Changes in tip clearance during a notional mission profile

During a notional mission profile, as shown in Figure 2, changing engine operating conditions result in great variation of blade-tip clearance, "pinch points" are created for large-magnitude transient events such as take-off and re-burst

(Decastro et al. 2004), and severe rubs will be caused between blades and shroud and it is dangerous for flight. Harsh working environments result in large radial deformation of disc, which is several times of blades deformation (Kypuros and Melcher 2003) and it is of paramount importance to these rubs.

Key loads on disc are: 1) High temperature/ pressure of mainstream flow; 2) Heat conduction between disc and blades; 3) Cooling effect by air along the borders; 4) Pulling force by rotating blades; 5) Centrifugal force generated by the rotating disc. To reduce the difficulty of optimization, other loads such as gravitational, gyroscopic and aero-dynamic loads are neglected. 4 types of BCs used are listed as follows:

1）Pressure: this is an equivalent load on rim by mainstream flow and pulling force by rotating blades, it is:

$$\tilde{P} = P_{gas} - n_b * \frac{m_b R_b \omega^2}{L * 2\pi * r} \qquad (1)$$

Where P_{gas} is pressure of mainstream flow, n_b is number of blades, m_b is mass of a blade, R_b is distance from blade centroid to disc axis, r is rim radius, ω is angular speed of rotor, L is rim thickness.

2）Body forces: Centrifugal force and gravity are applied to the disc.

3）Heat transfer with mainstream flow, could be written as:

$$-k \frac{\partial T}{\partial n}\bigg|_{\partial \emptyset} = \alpha h(T - T_f) \qquad (2)$$
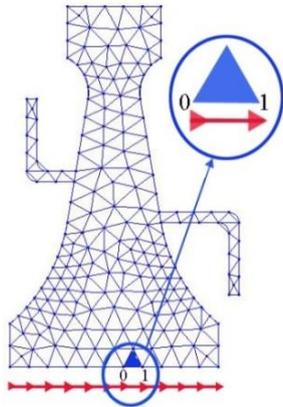

Figure: 3 Approximation BC of cooling air

Where k is thermal conductivity of disc, T is temperature of rim, T_f is temperature of mainstream flow, h is heat transfer coefficient and α is an experiential correction factor for heat from blades.

4）Convection with cooling air: Hot disc is cooled by air from compressor. Figure 3 gives an approximation for this fluid-thermal interaction, For boundary edge labeled by node 0 and 1, corresponding numerical equation is:

$$m_f * C * (T_{f1} - T_{f0}) = h * A * (T_f - T_w) \qquad (3)$$

Where $T_f = 0.5 * (T_{f1} + T_{f0})$, $T_w = 0.5 * (T_{w1} + T_{w0})$, T_{w0} and T_{w1} are temperatures of 2 nodes, T_{f0} and T_{f1} are temperatures of adjacent flow nodes, A is length of edge, m_f is mass of this air segment whose specific heat is C.

2. CONTRIBUTION RATIO BASED MAIN FACTORS SELECTION

Rough analysis shows that mass of web plate is manly decided by W3, W4, R3 and R4; mass of hub is decided by W2 and H2. Radial deformation is influenced by cooling and rotating factors, height of right flange (H6-H5) affects the cooling effect of right side, to the left flange, cooling effect is decided by H4, flanges connect disc with other components and they influence the radial deformation too. So (H1,H2,H4,H6,W2,W3,W4,W6,R3,R4) are set as factors of DOE whose responses are Mass and radial deformation. To simplify the computation, both flanges are set as same width (W6=W8), (W1,W5,W7,H3,H5,R1,R2,R5,R6) are set as fixed values; span and levels of each factor could be chosen with historical knowledge, Finally a Taguchi DOE with M samples is generated and main factors are selected as follows:

Firstly, each response is normalized by Eq. (4)：

$$b'_{ij} = \frac{b_{ij} - \min_{0 \le i \le M} b_{ij}}{\max_{0 \le i \le M} b_{ij} - \min_{0 \le i \le M} b_{ij}} \qquad (4)$$

Where b_{ij} is j-th response of i-th sample and entropy of j-th response is：

$$e_j = -q \sum_{i=1}^{M} p_{ij} \ln(p_{ij});$$

$$p_{ij} = \frac{b'_{ij}}{\sum_{i=1}^{M} b'_{ij}}; q = \frac{1}{\ln M} , j = 1,2 \qquad (5)$$

Weight of each response is: $w_j = \frac{\alpha_j}{\sum_{j=1}^{2} \alpha_j}$ (6)

Where $\alpha_j = \frac{1-e_j}{2-\sum_{i=1}^{2} e_i} , j = 1,2$

Secondly, range analysis is processed with the following steps:

1）Range of i-th factor on j-th response, denoted as R_{ij}, is formulated as:

$$R_{ij} = Max\left(\overline{K}_{(1,i),j} \dots \overline{K}_{(k,i),j}\right) - Min\left(\overline{K}_{(1,i),j} \dots \overline{K}_{(k,i),j}\right) \quad (7)$$

Where k is level number of i-th factor, $\overline{K}_{(l,i),j}$ is mean value of j-th response for l-th level of i-th factor.

2）Relative range of R_{ij} is: $\quad R'_{ij} = \frac{R_{ij}}{\sum_{i=1}^{N} R_{ij}} \quad (8)$

3）Comprehensive relative range of i-th factor is:

$$R'_i = \sum_{j=1}^{n} w_j R'_{ij} \quad (9)$$

Finally, factors are sorted by R'_i, and "contribution ratio" is used to select main factors for ANN. For certain disc, the corresponding cumulative contribution ratios are plotted in Figure 4, cumulative contribution ratio of first 8 factors is 0.9628, and corresponding variables are (W6,H1,W3,W4,R4,R3,W2,H6).
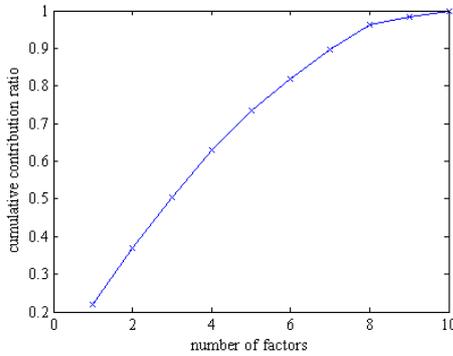

Figure 4: Variation of cumulative contribution ratio with number of factors

3. BPNN training with virtual samples

Main factors selected above and responses of DOE are respectively used as inputs and outputs of NN. Suppose relation between input vector $x \in R^{m \times 1}$ and output vector $y \in R^{n \times 1}$ is:

$$y = f(x) \quad (10)$$

A three-layered BPNN with topology "m-p-n" is used for approximation:

$$\tilde{y} = \tilde{f}(\omega, x) \quad (11)$$

Optimized ω with weight number $(m + 1)p + (p + 1)n$ should be found to minimize estimation error. Levenberg-Marquardt algorithm, an enhanced method of Moore-Penrose for over-determined question, is used to train NN in this paper. If ratio ρ is used to select training samples for cross-validation, size of available data set,

denoted as M', should be:

$$\rho M' \geq (m + 1)p + (p + 1)n \quad (12)$$

Normally, size of DOE samples, denoted as M, is much less than M' and virtual samples could be used for this M'. Define i-th DOE sample as $S_i = \{x_{i1} \dots x_{ij} \dots x_{i(m+n)}\}$, where $x_{i(m+j)} = y_{ij}$ and $x_{ij} \in [-1,1]$. For each x_{ij}, the perturbation result is:

$$x'_{ij}$$

$$= \begin{cases} 2 - (x_{ij} + \beta * d_{ij}) & if \quad x_{ij} + \beta * d_{ij} \geq 1 \\ x_{ij} + \beta * d_{ij} & else \\ -2 - (x_{ij} + \beta * d_{ij}) & if \quad x_{ij} + \beta * d_{ij} \leq -1 \end{cases}$$

Where $d_{ij} \in [-1,1]$ is random noise and $\beta \in [0,0.5]$ is noise intensity. This is a mirroring operation with threshold -1 and 1.

If virtual samples are used to train BPNN, Eq. (11) will be:

$$\tilde{y} + \Delta y = \tilde{f}(\omega, x + \Delta x) \quad (14)$$

Error of BPNN from Eq. (14) is:

total error= approximation error
+ generalization(estimation) error
+ sample error

Approximation error is decided by training algorithm; generalization error is decided by BPNN structure and coverage of problem domain; sample error is decided by size of virtual sample and noise intensity. If size of virtual samples is denoted as M_v, total error will be formulated as:

$$\varepsilon^* = G(m, n, \beta, p, M, M_v) \quad (15)$$

$M_v = PM$ is used to generate uniformly distributed virtual samples where P is an integer. With Eq. (12), p will be:

$$p \leq \left\lfloor \frac{\rho * M * (P + 1) - n}{m + n + 1} \right\rfloor \quad (16)$$

For given P, the maximum p will be the best choice with the consideration of calculation complexity and generalization ability, and Eq. (15) can be simplified as:

In Vapnik's theoretical framework, approximation error increases with the increase of VC dimension (size of hidden layer, p), but estimation error decreases with the increase of p (Haykin 2011), the optimum p is determined by the point at which both errors assume a common value. Before the optimum condition is reached, the learning problem is over determined, beyond the minimum point, the learning problem is underdetermined. So a good p (and corresponding P) should be found to match size of DOE samples. Energy of sample error for this P is:

$$\varepsilon_{sample} = \frac{1}{2} \sum_{i=1}^{PM} \sum_{j=1}^{j=m+n} \beta^2 \varepsilon_{d_{ij}} \quad (18)$$

Where $\varepsilon_{d_{ij}}$ is energy of noise d_{ij} that can be denoted

asε_d, Eq. (18) is simplified as:

$$\varepsilon_{sample} \approx \frac{1}{2} PM(m+n)\beta^2 \varepsilon_d \qquad (19)$$

d_{ij} is a random value, so there is a suitable "sample error" for given DOE samples, and it is reasonable to find an optimized pair (P, β) for this "sample error".

Strong noise disturbs original regularities which increases "approximation error", weak noise is similar to boosting which increases "generalization error", experiments show that for given P, there is a best β to generate virtual samples.

4. NSGA-II for HPTD

Evolutionary algorithms can search multiple objectives in parallel, among which NSGA-II is a good choice because it accelerates the optimization speed with an efficient non dominated sorting method; quality of the population is improved with crowding algorithm and elitist strategy instead of sharing function (Deb 2002). However, its computing ability is limited by population and iteration of GA, a good initial population should be used to improve the efficiency and quality of selection and evolution.

Normally, initial population of GA is randomly generated for diversification and a large population is used to reach a global convergence and accuracy, but the optimization process will be prolonged significantly for complex problems. Some researchers take DOE samples and random samples as initial population. This paper constructs the initial population with 3 subsets, i.e. DOE samples, virtual samples and random samples, ratio for them are defined tentatively (for example 2:2:1). In Figure 5, advantage of this synthetic method is described by the comparing of four Pareto fronts resulting from different initial populations. Obviously, Pareto front evolved from random samples is much more uniform (Figur. 5a, 5c), front evolved from DOE samples is much more smooth (Figure 5b), front evolved from a combination of DOE samples and virtual samples takes both of these advantages (Figure 5d).

(a) Pareto front from 36 random samples   (b) Pareto front from 36 DOE samples

(c) Pareto front from 72 random samples   (d) Pareto front from 36 DOE & 36 virtual samples
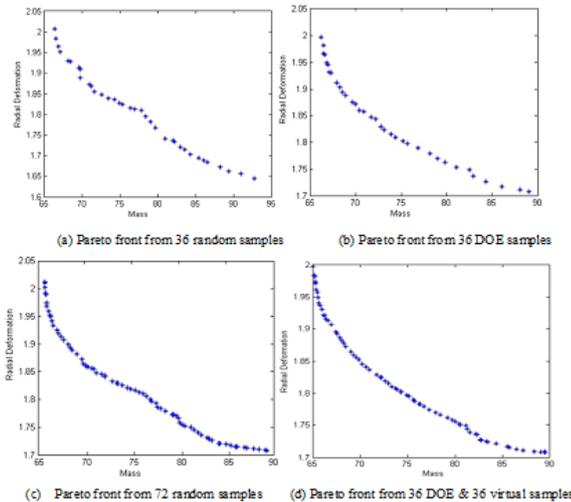
Figure 5: Pareto front based on different initial population

Designer wants only one or several optimal solutions instead of so many Pareto optimal outcomes. Several methods can be used to select better solutions from this front. In this paper, "comprehensive proportion" is used as criterion for selection which is described as follows:

1）For Pareto front with size *K*, each objective b_{ij} is renormalized as μ_i^j:

$$\mu_i^j = \frac{b_{ij} - \min_{0 \le i \le K} b_{ij}}{\max_{0 \le i \le K} b_{ij} - \min_{0 \le i \le K} b_{ij}} , \qquad (20)$$
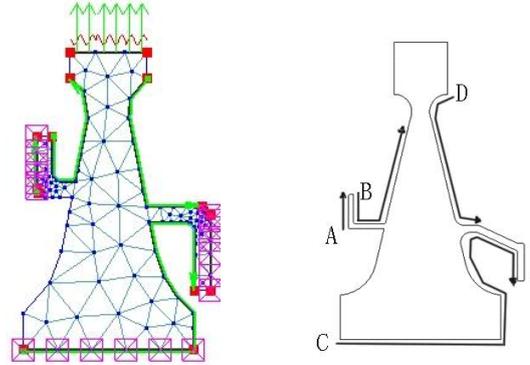$$i = 1,2 \dots K; \qquad j = 1,2 \dots n$$

2）Proportion of *i*-th solution on *j*-th objective is calculated with Eq. (21):

$$u_i^j = \frac{\mu_i^j}{\sum_{i=1}^K \mu_i^j} \qquad (21)$$

3）With weights from Eq. (6), a comprehensive proportion and a normalized value are calculated by Eq. (22):

$$u_i = \sum_{j=1}^n w_j u_i^j ; \qquad u_i' = \frac{u_i}{\sum_{i=1}^K u_i} \qquad (22)$$

4）Good solutions are selected by u_i', The smaller u_i' is, the better a Pareto optimal will be.

(a)Coupled BCs   (b) Distribution and direction of cooling flows

Figure 6: details of thermo-structure BCs for "take-off"

4. Case study

In this paper, targets of optimization are: 1) Minimum mass (*W*) that should be less than 80kg; 2) Minimum radial deformation (*D*) that should be less than 1.95mm. 68 blades are fixed into mortises, each blade's mass is 0.144kg and

height is 56mm, distance from blade centroid to rotating axis is 298mm. Coupled flow-thermo-structure BCs for "take-off" are shown in Figure 6a where "↑" represents equivalent pressure, "~" indicates heat transfer with mainstream flow, "⊠" denotes zero displacement, 4 curves along border are cooling flows whose details are described by Figure 6b, parameters of this state are specified in Table 1.

Table 1 Working parameters for "take-off"

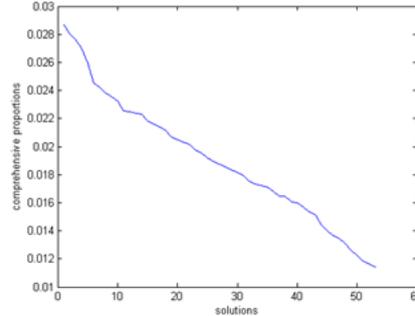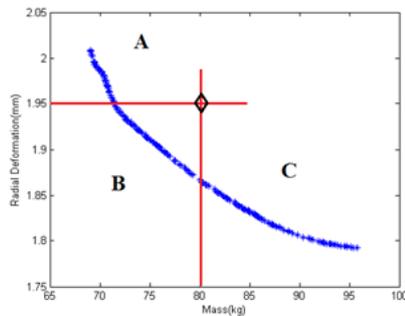| Variable	Value	Variable	Value
Disc density	8210kg/m3	rotation speed of disc	12,500 rpm
Gas temperature	1650k	convection coefficient	2200w/m2.k
Pressure in duct	35 atm		
speed of flow A	0.1918kg/s	source temperature of A	907k
speed of flow B	2.4754kg/s	source temperature of B	968k
speed of flow C	0.5139kg/s	source temperature of C	681k
speed of flow D	0.7666kg/s	source temperature of D	791k

(H1,H2,H4,H6,W2,W3,W4,W6,R3,R4) are selected as factors for DOE, span of each factor is summarized in Table 2, and (W1,W5,W7,H3,H5,R1,R2,R5,R6) are set as (45.0,42.0,60.0,195.0,140.0,280.0,105.0,5.0,6.0).

Table 2 Spans of 10 factors for certain turbine disc

(unit:mm)

NO	factor	minimum	maximum	NO	factor	minimum	maximum
1	W6	5.0	9.0	6	W3	15.0	25.0
2	H1	15.0	25.0	7	W4	45.0	55.0
3	H4	30.0	40.0	8	H2	14.0	28.0
4	H6	165.	175.	9	R3	235.	247.
5	W2	90.0	110.	10	R4	155.	175.

With the help of historical knowledge, A mixed-level design is generated with (W2,W3,W4,H2,R3,R4) at three levels and (W6,H1,H4,H6) at two levels, DOE of L36(3^6*2^4) is constructed.

By sensitivity analysis, variables (W6,H1,W3,W4,R4,R3,W2,H6) are used as inputs of BPNN whose topology is "8-p-2", different (P,β) pairs are used to generate virtual samples. Levenberg-Marquardt algorithm with goal=5e-6 and ρ=0.9 is used to train BPNN. for given P, p should be:

$$p \leq \left\lfloor \frac{\rho * M_1 * (P+1) - n}{m+n+1} \right\rfloor = \left\lfloor \frac{0.9 * 36 * (P+1) - 2}{11} \right\rfloor$$

Good (P,p) pairs are (1,5),(2,8),(3,11),(4,14),(5,17) and so on, several training sets are created for each pair with a series of βs whose initial value is 0.1 with step size $\sqrt{0.1}$ (i.e., $\beta = 0.1,0.0316,0.01,0.00316\ldots$), best error for each "8-p-2" BPNN and corresponding β are list in Table 3.

Table 3 36 test error of BPNNs with different topologies

ID	(P,p)	β	test error	ID	(P,p)	β	test error
1	(1,5)	0.0316	1.13e-2	4	(4,14)	0.001	1.01e-2
2	(2,8)	0.00316	7.21e-3	5	(5,17)	0.00036	1.20e-2
3	(3,11)	0.00316	9.85e-3	6	(6,20)	0.0001	1.74e-2

BPNN with structure "8-8-2" is selected as fitness function for further NSGA-II whose parameters are defined in Table 4. The initial population consists of DOE samples, virtual samples and random samples. Ratio of them is 63:63:32.

Table 4 Specifications of MOO

Parameter	Value	Parameter	Value
population	158	decisions/objectives	8/2
Generation	1000	Selection and replacement	NSGA-II
crossover probability	0.9	Mutation probability	0.1
crossover operator	Simulated binary	Mutation operator	Polynomial
distribution index for crossover	20	distribution index for Mutation	20

Figure 7: Result of NSGA-II

The scatter plot of the two objectives by Pareto front is illustrated in Figure 7a, 53 solutions in area B are useable points, and location of "◊" is decided by optimization targets.

Normalized comprehensive proportions of these 53 solutions are calculated and sorted in Figure 7b, 5 minimum solutions are better than the others. FEM of these 5 solutions are reanalyzed and the best one is selected as the optimization solution. It can be noted that for given optimization targets, W6 should be set as maximum value to minimize radial deformation, H1,W2,W3,R3,R4 could be set as minimum value to minimize the mass,W4 and H6 should be designed carefully. That is to say, flange details influence radio deformation greatly.

5. CONCLUSIONS

In this paper, mass and radial deformation are taken as mutually restricted targets to design HPTD. Virtual samples are generated to enlarge training set of NN; the generalization performance of NN, which is very important for MOO, is improved by the compromise of "hidden layer size", "virtual sample size" and "noise intensity"; the smoothness and uniformity of Pareto front are improved by the hybrid initial population who is a combination of DOE samples, virtual samples and random samples. Latter technologies are explored based on former ones; MOO of HPTD is carried out step by step.

Noise intensity (β) for virtual samples will be optimized in continuous space for a global optimal. More work will be done to compress the redundancy of DOE samples, metric such as max-min distance will be applied on these samples to select a subset with high information density, the computational cost will be reduced and advantage of virtual samples will be fully exploited. Furthermore, other surrogates will be studied on the virtual samples to provide fitness functions with better generalization ability.

ACKNOWLEDGMENTS

REFERENCES

Bakhtiari, H.M. Karimi, S. Rezazadeh. (2016) Modeling, analysis and multi-objective optimization of twist extrusion process using predictive models and meta-heuristic approaches, based on finite element results. *Journal of Intelligent Manufacturing* 27(2):463-473

Bishop, C.M. (1995) Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, 7(1): 108-116.

Deb, K., A. Pratap, S. Agarwal and T. Meyarivan. (2002) A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2): 182-197.

Decastro, J.A. and K.J. Melcher. (2004) A Study on the requirements for fast active turbine tip clearance control systems. *NASA /TM-2004-213121*.

Hailia, E.E., D.T. Lenahan and T.T. Thomas. (1982) High pressure turbine test hardware detailed design report. *NASA CR-167955*.

Haykin, S. (2011) *Neural Networks and Learning Machines (3rd Edition)*. New Jersey: Pearson Prentice Hall

Jin, Y.C. (2011) Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation* 1(2):61-70

Joly, M.M., T. Verstraete and G. Paniagua (2013) Differential evolution based soft optimization to attenuate vane–rotor shock interaction in high-pressure turbines. *Applied Soft Computing* 13(4): 1882–1891.

Jahed, H., B. Farshi and J. Bidabadi.(2005) Minimum weight design of inhomogeneous rotating discs. *International Journal of Pressure Vessels and Piping* 82:35-41.

Kypuros, J.A. and K.J. Melcher. (2003) A reduced model for Prediction of Thermal and Rotational effects on Turbine Tip Clearance. *NASA/TM—2003-212226*.

Li, D.C., W.C.Chen, C.J. Chang, C.C. Chen and I.H. Wen. (2015) Practical information diffusion techniques to accelerate new product pilot runs. *International Journal of Production Research* 53 (17): 5310-5319.

Li, H., W. Water, B. Zhu, J. Lu. (2015) Integrated high frequency coaxial transformer design platform using artificial neural network optimization and FEM simulation. *IEEE Transactions on Magnetics* 51 (3):1-4

Nagendra, S., J.B. Staubach, A.J. Suydam,et al (2005) Optimal rapid multidisciplinary response networks: RAPIDDISK. *Structural and Multidisciplinary Optimization* 29(3):213-231

Niyogi, P., F. Girosi and T. Poggio.(1998) Incorporating prior information in machine learning by creating virtual examples. *Proceedings of the IEEE*, 86(11):2196-2209.