Supplemental material for Zhao et al., "An Improved Water Wave Optimization Algorithm with the Single Wave Mechanism for No-wait Flow Shop Scheduling Problem," *Engineering Optimization*, 2018.

S1. Computational Complexity for SWWO

SWWO is made up of four parts. The first part is to initialize population and a new initialization sequence method is proposed, named NEH_COV, the time complexity of initialization is $O(n^2)$. The second part is the propagation operation. Based on the block-shift operation, the modified BSO, named MBSO, is proposed. The time complexity of the second part is $O(n^2)$. The third part is the refraction operation. There is only one water wave in SWWO. A new solution is obtained by crossing the historical optimal solution with the current solution. The time complexity is $O(mn)$. The fourth part is the breaking operation. By executing an improved variable neighborhood search, named MVNS, the time complexity is $O(2n^2)$. According to the analyses above, the time complexity of SWWO under the condition of k times iteration is calculated as follows.

$$\begin{aligned} O(k,m,n) &= O(n^2) + O(k) + [O(n^2) + O(mn) + O(2n^2)] \\ &\approx O(n^2) + O(k) \times O(n^2) \\ &\approx O(kn^2) \end{aligned}$$

S2. Evaluation Criterion

In order to demonstrate the effectiveness of our proposed algorithm and compare the performance with other approaches visually, the average relative percentage deviation (ARPD) was applied to measure the quality of the experimental results. The calculation formula of ARPD is as follows:

$$ARPD = \frac{1}{R}\sum_{r=1}^{R}\frac{C_r - C_R^*}{C_R^*}\times 100. \quad (1)$$

where C_r is the solution of the rth experiment, which is generated by a specific algorithm, and C_R^* is the optimum solution found so far. In addition, the standard deviation (SD) is also recorded to indicate the robustness of the algorithm, where the SD is calculated as follows:

$$SD = \sqrt{\frac{\sum_{r=1}^{R}(C_r - \overline{C})^2}{R}}. \quad (2)$$

where C_r is the solution of the rth experiment, which is generated by a specific algorithm, and $\overline{C}$ is the mean value of R solutions.

S3. Parameters analysis

The SWWO has three parameters, λ_{max}, h_{max}, T_0. The Design of Experiments (DOE) (Montgomery 2001) approach is used for the parameter tuning of the algorithm. More precisely, the total factor analysis of the three parameters as factor is carried out. The parameters are as follows.

- λ_{max}, 5 levels, 6,8,10,12,14
- h_{max}, 6 levels, 6,8,10,12,14,16
- T_0, 6 levels, 0,2,4,6,8,10

The different values of three parameters produce 180 different combinations of parameters. To avoid over-fitting the parameters, the program using different data sets is executed during the parameter setting phase and experimental evaluation phase. In parameter setting phase, the experiments are performed on 21 sets of VRF instances (Vallada, Ruiz, and Framinan 2015), which include {20,40,60}×{5,10,15,20}, {100,300,500}×{20,40,60}. In each combination, 10 repeat trials are run and ARPD values are recorded for each instance.

The experiment is analyzed by means of a multi-factor Analysis of Variance (ANOVA) technique. To apply ANOVA, the three main hypothesis of ANOVA, i.e., normality of data, homoscedasticity and independence of residuals are checked. The residuals resulting from the experimental data are analyzed and three hypotheses are accepted. The ANOVA results are shown in Table 1. All the three parameters and their interactions have small *p*-value. When the *p*-value is close to zero, *F*-ratio replaced the function of *p*-value to show the significance of parameters. A large *F*-ratio has a considerable effect on the algorithm. From the result of *F*-ratio, the influence of the parameter $\lambda_{\max}$, $h_{\max}$, T_0 is very significant, and their interaction has a relatively small influence.

Table 1. ANOVA results for the experiment on tuning parameters

Source	Sum of squares	D_f	Mean square	*F*-Ratio	*p*-Value
Main effects					
A:λ_{max}	0.074	4	0.019	547.485	0.000
B:h_{max}	0.007	5	0.001	41.388	0.000
C:T_0	0.042	5	0.008	247.695	0.000
Interaction					
AB	0.003	20	0.000	3.866	0.000
AC	0.004	20	0.000	6.451	0.000
BC	0.005	25	0.000	5.460	0.000
ABC	0.005	100	5.188E-5	1.527	0.003
Residual	0.012	360	3.398E-5		
Total(Corrected)	0.153	539			

Table 2. Multiple Comparison of parameters

Parameter	levels	Average ARPD	95% Confidence interval	99% Confidence interval
λ_{max}	14	0.134	a	A
	6	0.113	b	B
	12	0.107	c	C
	10	0.105	c	C
	8	0.102	d	D
h_{max}	6	0.119	a	A
	8	0.114	b	B
	14	0.110	c	C
	12	0.110	c	C
	16	0.110	c	C
	10	0.109	c	C
T_0	0	0.131	a	A
	2	0.112	b	B
	4	0.108	c	C
	8	0.107	c	C
	10	0.107	c	C
	6	0.106	c	C

In order to make a significant analysis of the levels of the parameters and determine the values of the parameters, a Tukey test is applied to examine the differences between levels of parameters. From Table 2, in the case of 95% and 99% confidence interval, the levels of the parameter λ_{max} are divided into four subsets, {14}, {6}, {10, 12}, {8}. Among the subsets, the levels of the parameter are significantly different. There is no significant difference between the levels of the parameter within the subset. It is clear from Figure 1 that too small or too large value of λ_{max} leads to deterioration of the performance of algorithm and the choice of $\lambda_{max} = 8$ gives the best result. For parameter h_{max}, in the case of 95% and 99% confidence interval, the levels are divided into three subsets, {6}, {8}, {10, 12, 14, 16}. Confidence interval suggests that the algorithm is robust for the h_{max} in the range [10, 16]. For parameter T_0, in the case of 95% and 99% confidence interval, the levels are divided into three subsets, {0}, {2}, {4, 6, 8, 10}. From Figure 1, the ARPD obtained by running the algorithms at T_0 = 0 are worse than other non-zero values.

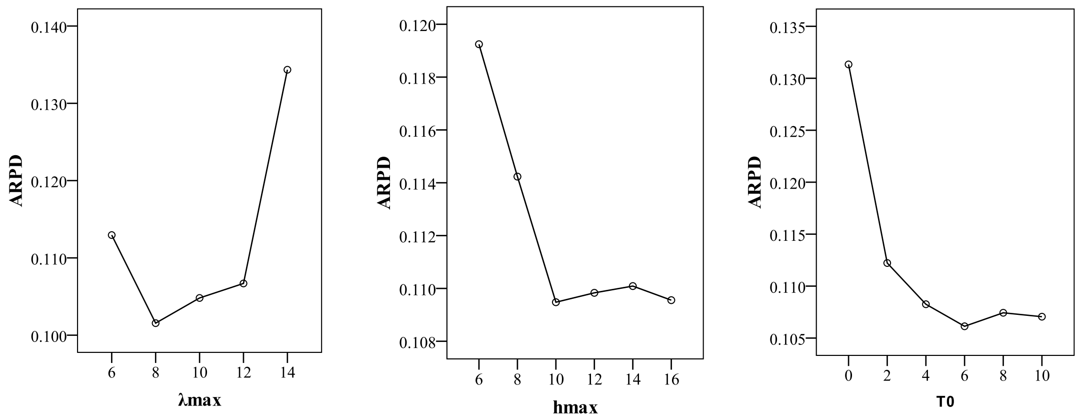

Figure 1. Factor level trend of SWWO

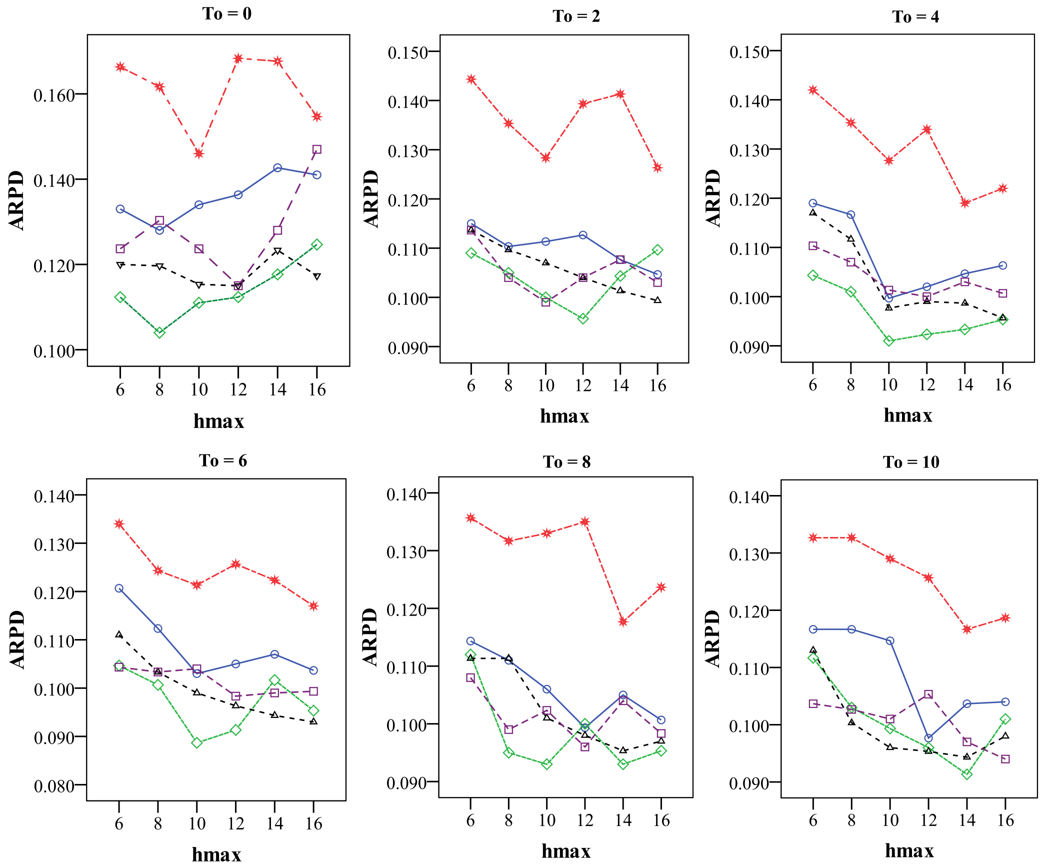

Figure 2. Interaction plots of factor levels

However, the main effect is meaningless if there is a significant interaction between the parameters. From Table 1, the interaction $\lambda_{max}\times h_{max}\times T_0$ is significant because the p-value is less than $\alpha = 0.05$. It is necessary to check the interaction plot whether it deviates from the judgments in the main effects plot. The interaction plot is given in Figure 2. The results are in accordance with the conclusions. From the results of the experimental analysis, the final parameter settings utilized in SWWO are as follows, $\lambda_{max} = 8$, $h_{max} = 10$, $T_0 = 6$.

S4. Analysis of neighborhood structures

The neighborhood structures selected and order used in search process affect the efficiency of the performance of VNS. In this paper, two neighborhood structures are applied to breaking wave. Since the propagation operation is based on the extension of the basic insertion operation, the breaking operation is performed firstly in the exchange neighborhood structure of the sequence and then in the basic insertion neighborhood structure. The refraction operation is a desirable feature for learning superior solutions. In the combinatorial optimization problem, insertion and swap operators are a basic and efficient local search strategy. Firstly, is performed in the swap neighborhood structure of the sequence. Secondly, the breaking operation is performed in the basic insert neighborhood structure. Therefore, swap operation and insert operation greatly improve the local search ability and find the optimal solution.

References

Montgomery, D. C. 2001. *Design and analysis of experiment*. Phoenix, AZ: John Wiley and Sons.

Vallada, Eva, Rubén Ruiz, and Jose M. Framinan. 2015. "New hard benchmark for flowshop scheduling problems minimising makespan." *European Journal of Operational Research* 240 (3): 666-677.