Supplementary material

1. Taguchi design

Taguchi aims to determine the settings of the GA-VNS parameters so that at those levels, the algorithm is made insensitive to variations in the noise factors without actually eliminating them (Nourmohammadi and Zandieh 2011). Thus, by applying the Taguchi experiments over different test problems and their relating capacities considered as the noise factors, the best GA-VNS parameters' levels are chosen so that the resulting S/N ratios and the TC means, are maximized and minimized, respectively (Taguchi et al. 2007). Figure 1 shows the means and the S/N ratios diagrams over different combinations of GA-VNS parameters' levels.

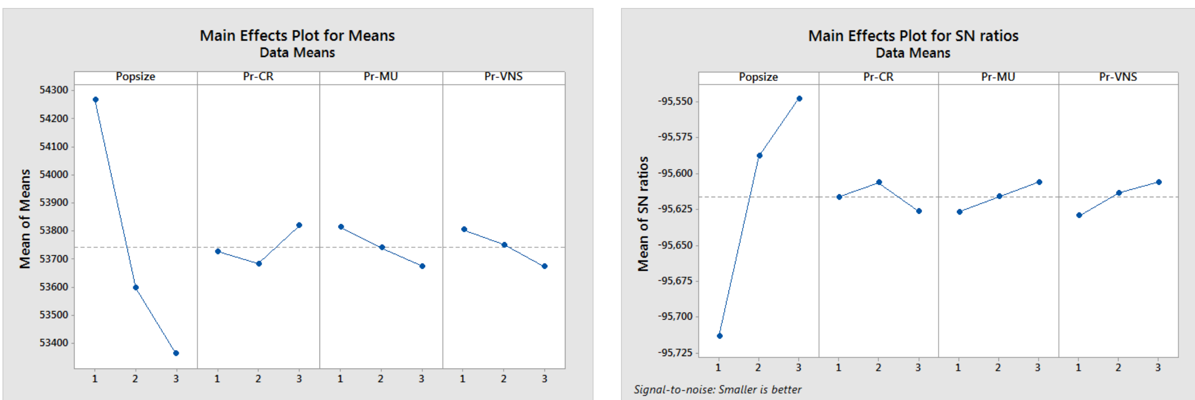
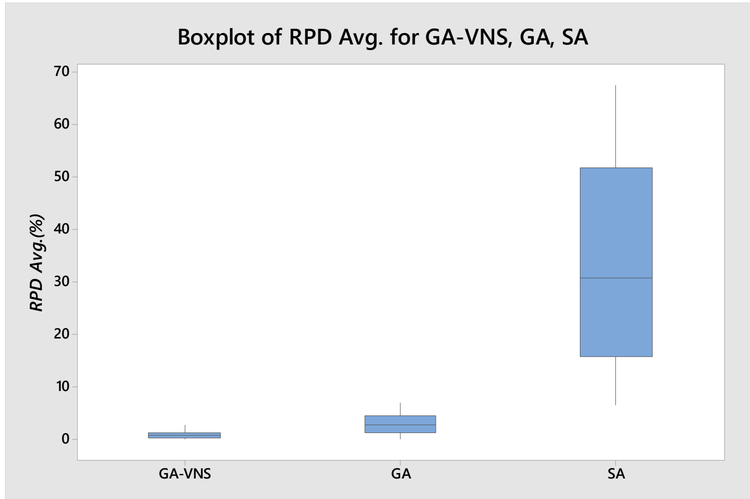


Figure 1. The means (the left) and the S/N ratios (the right) diagrams over different GA-VNS parameters' levels.

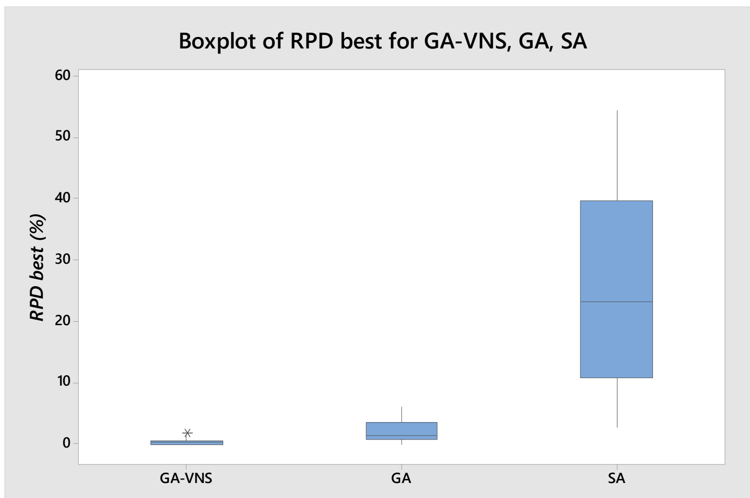
It is worth noting that although in the parameter tuning process all the problems instances are taken into considerations, the results can only guarantee the robustness of GA-VNS, which is stochastic by nature, over the considered test problems.

2. Statistical analysis of the results

To conduct a valid comparison for assessment of the algorithms' performance, the boxplots for the RPD_{Avg} and RPD_{best} of the compared meta-heuristics are plotted in Figure 2(a) and 2(b), respectively.



(a)



(b)

Figure 2. Boxplots of (a) RPD_{Avg} and (b) RPD_{best} for GA-VNS, GA and SA.

According to Figure 2, one can observe that GA-VNS outperforms GA and SA in terms of both RPD_{Avg} and RPD_{best}. However, in order to obtain rigorous and fair conclusions, a statistical test is performed to statistically validate the hypothesis whether there are significant differences among the performances of the meta-heuristics in terms of the RPD_{Avg} and RPD_{best}. Thus, nonparametric Friedman F_r test for a randomized block design is applied to determine whether evidence exists to indicate that the RPD_{Avg} and RPD_{best} differ in location for at least two algorithms. The null and alternative hypotheses are H_0: the probability distribution of the RPD_{Avg} or RPD_{best} of the algorithms are identical for three algorithms and H_1: the probability distribution of the RPD_{Avg} or RPD_{best} of at least two algorithms are significantly different from one another. The resulting Friedman F_r statistics for the RPD_{Avg} and RPD_{best} have been 72.47 and 73.23, respectively. Assuming that the confidence level of the test is 99% (α=0.01), the critical point in a χ^2 distribution with 2 degrees of freedom is 9.21. Since both 72.47 and 73.23 are greater than 9.21, it can be concluded that there are significant differences in the RPD_{Avg} and $RPD_{best}$of at least two compared algorithms, being GA-VNS the one with the lowest ranks for both of the RPD_{Avg} and RPD_{best} as shown in Table 1.

Table 1. Friedman statistical test on the $RPD_{Avg.}$ and RPD_{best}.

Algorithm	No. of samples	$RPD_{\text{Avg.}}$		RPD_{best}	
		Median	Sum of Ranks	Median	Sum of Ranks
GA-VNS	40	0.75	44.5	0.22	46.5
GA	40	2.86	75.5	1.71	73.5
SA	40	30.43	120	22.14	120
Overall	120	11.35		8.02	

To evaluate the statistical significance of the better performance of GA-VNS, the Holm's post hoc test has also been conducted in which GA-VNS is considered as the control algorithm. Using this test a paired comparison between GA-VNS and other algorithms is performed and the resulting unadjusted and adjusted p-values obtained for the RPD_{Avg} and RPD_{best} can be seen in Table 2.

Table 2. Unadjusted and adjusted p-values for the $RPD_{Avg.}$ and RPD_{best} using Holm's post hoc test by setting GA-VNS as the control algorithm.

Comparison	$RPD_{\text{Avg.}}$		RPD_{best}	
	Un Adj. p-value	Adj. p-value	Un Adj. p-value	Adj. p-value
GA-VNS vs SA	0	0	0	0
GA-VNS vs GA	0.000264	0.000132	0.00127	0.000635

Since all the resulting p-values are lower than 0.01, it can be inferred that GA-VNS is significantly better than both SA and GA in terms of the RPD_{Avg} and RPD_{best} at a 99% confidence level.