Automatically Improving the Anytime Behaviour of Optimisation Algorithms: Supplementary material☆

Manuel López-Ibáñez*, Thomas Stützle

IRIDIA, Université Libre de Bruxelles (ULB), CP 194/6, Av. F. Roosevelt 50 – B-1050 Brussels, Belgium

Abstract

Optimisation algorithms with good anytime behaviour try to return as high-quality solutions as possible independently of the computation time allowed. Designing algorithms with good anytime behaviour is a difficult task, because performance is often evaluated subjectively, by plotting the trade-off curve between computation time and solution quality. Yet, the trade-off curve may be modelled also as a set of mutually nondominated, bi-objective points. Using this model, we propose to combine an automatic configuration tool and the hypervolume measure, which assigns a single quality measure to a nondominated set. Our goal is to improve the anytime behaviour of optimisation algorithms by means of automatically finding algorithmic configurations that produce the best nondominated sets. Moreover, the recently proposed weighted hypervolume measure is used here to incorporate the decision-maker's preferences into the automatic tuning procedure. We report on the improvements reached when applying the proposed method to two relevant scenarios: (*i*) the design of parameter variation strategies for MAX-MIN Ant System, and (*ii*) the tuning of the anytime behaviour of SCIP, an open-source mixed integer programming solver with more than 200 parameters.

Paper extended version	`http://iridia.ulb.ac.be/IridiaTrSeries/IridiaTr2012-012.pdf`
TSP instances	`http://iridia.ulb.ac.be/supp/IridiaSupp2012-011/`
SCIP instances	`http://www.cs.ubc.ca/labs/beta/Projects/ParamILS/results.html`

☆Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043

*Corresponding author

Email addresses: manuel.lopez-ibanez@ulb.ac.be (Manuel López-Ibáñez), stuetzle@ulb.ac.be (Thomas Stützle)

Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043

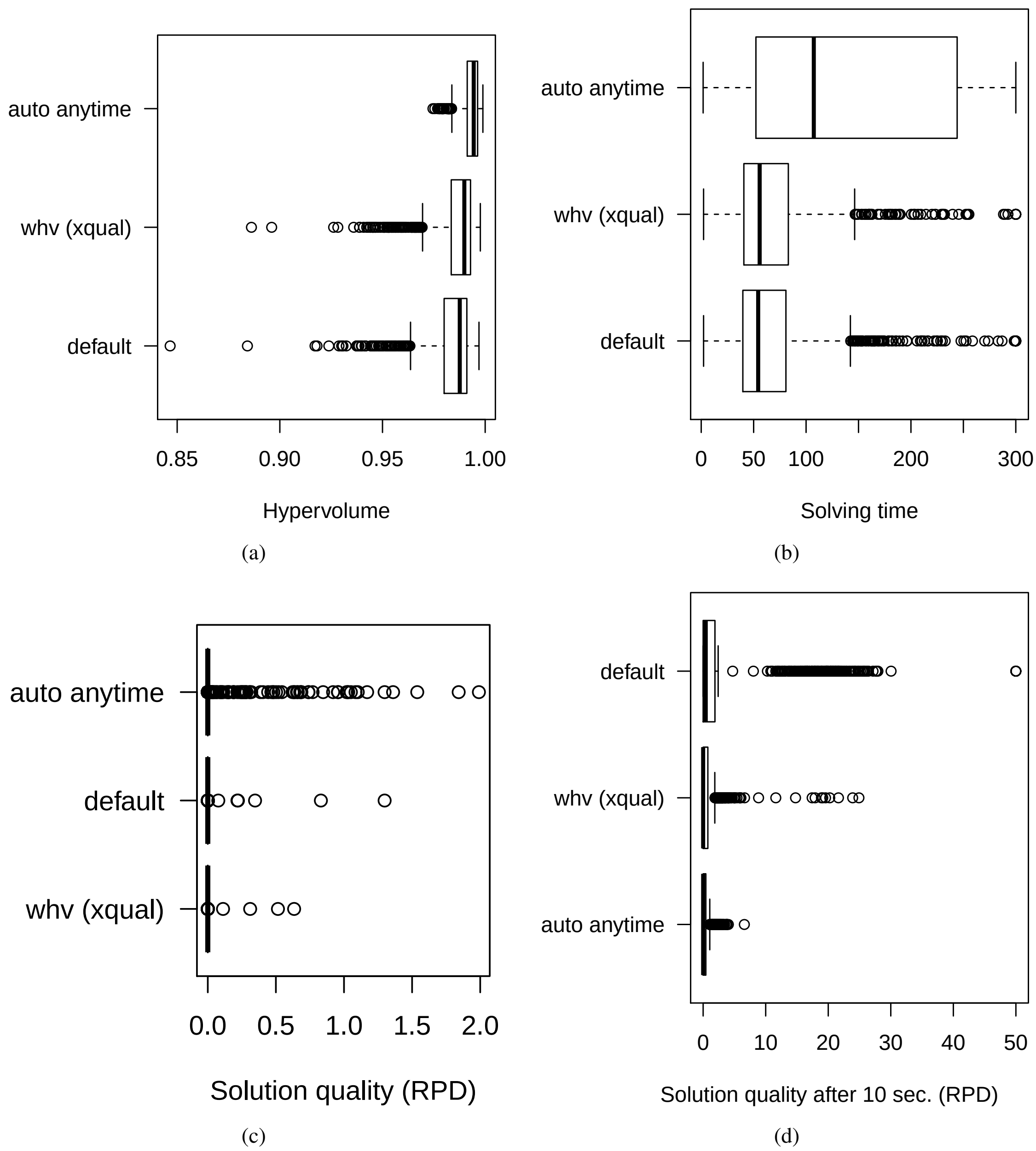


Figure 1: Boxplots of the results obtained by three different parameter configurations of SCIP according to various evaluation criteria.

Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043

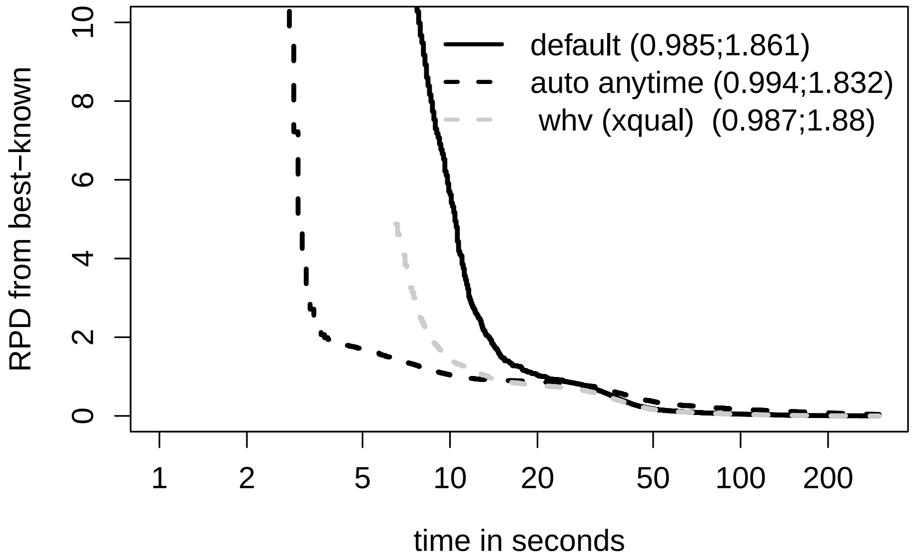


Figure 2: Mean RPD over all test instances for different configurations of SCIP. The number in parentheses is the mean hypervolume corresponding to that configuration.

Table 1: Configurations of SCIP ordered according to the sum of ranks with respect to three different evaluation criteria. The numbers in parenthesis are the difference of ranks relative to the best ranked configuration. Configurations that are not significantly different from the best one according to the Friedman test are indicated in bold face.

ΔR_α	**Configurations** (ΔR)	**Evaluation criterion**
		Time to best found
45.81	**default** (0), whv (xqual) (189.5), auto anytime (1558)	
		Quality after 10 seconds
77.88	**whv (xqual)** (0), auto anytime (296), default (488.5)	
		Final quality (300 seconds)
∞	**whv (xqual)** (0), **default** (19), **auto anytime** (54.5)	

Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043

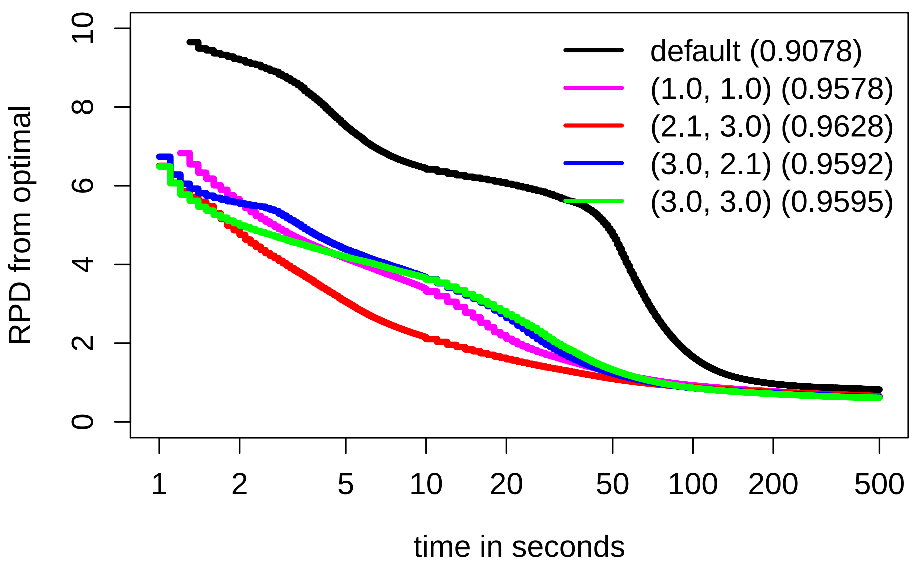


Figure 3: Anytime behaviour of the default configuration (default) vs. automatically tuned configurations of MMAS. Each line is the mean RPD over time averaged over 15 runs on each TSP instance in the test set. Automatically tuned configurations were obtained by using different normalizations and reference points during tuning. In particular, the configuration denoted by $(1.0, 1.0)$ used a normalization to the $[0.0, 0.9]$ interval, while the others used a normalization to the $[1, 2]$ interval. It is expected that a higher value in the second component of the reference point, such as in $(2.1, 3.0)$ will produce a configuration that prefers lower time for the same RPD rather than lower RPD for the same time. This would explain why the line corresponding to $(2.1, 3.0)$ converges faster but to a worse final RPD. The opposite is expected for $(3.0, 2.1)$, but the effect is not clearly visible, which may be due to a lower bound effect on the final quality that MMAS is able to achieve or simply to the stochastic nature of the algorithms. A more crucial observation is the significant improvement in anytime behaviour of all automatically tuned configurations, independently of the reference point used for the tuning, over the default configuration.

Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043

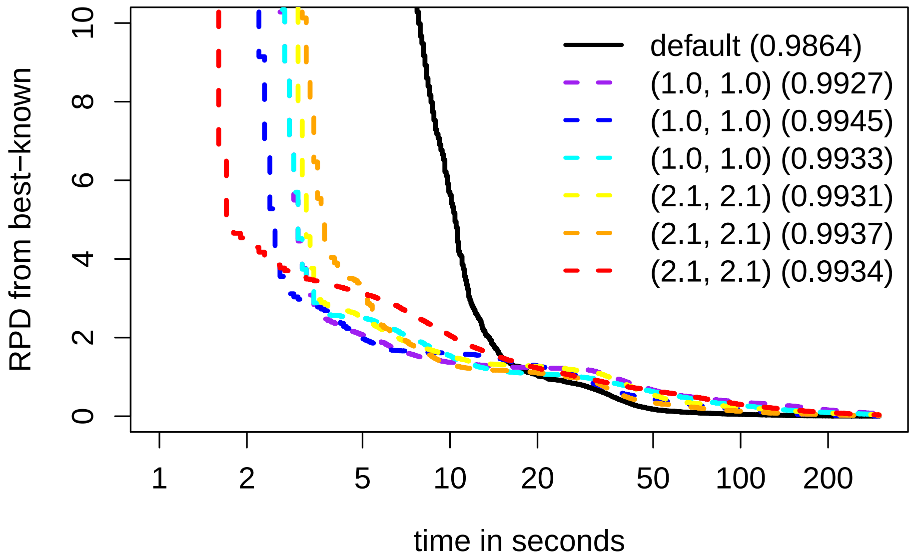
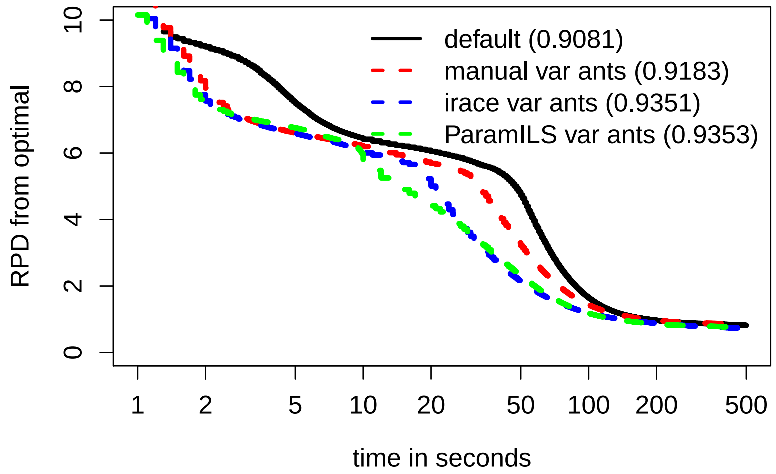
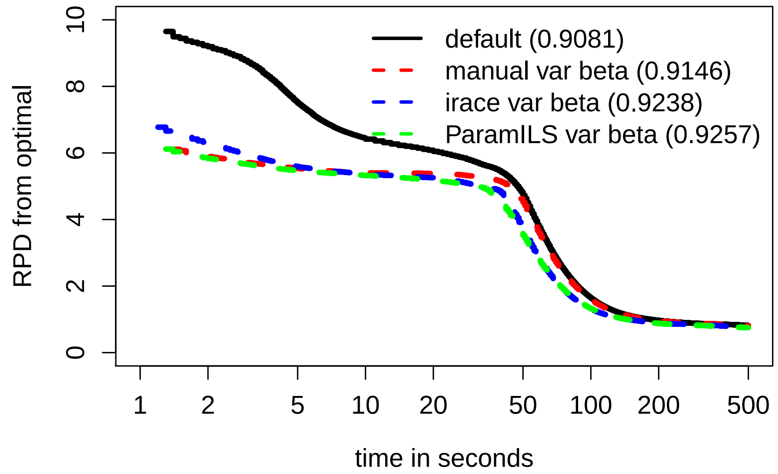


Figure 4: Mean RPD over all test instances for different configurations of SCIP. The number in parentheses is the mean hypervolume corresponding to that configuration computed using $[0.0, 0.9]$ as normalization bounds and $(1.0, 1.0)$ as reference point. Configurations denoted by $(2.1, 2.1)$ were tuned using the hypervolume with $[1, 2]$ as normalization bounds and $(2.1, 2.1)$ as reference point. Configurations denoted by using $(1.0, 1.0)$ were tuned using the hypervolume with $[0.0, 0.9]$ as normalization bounds and $(1.0, 1.0)$ as reference point.

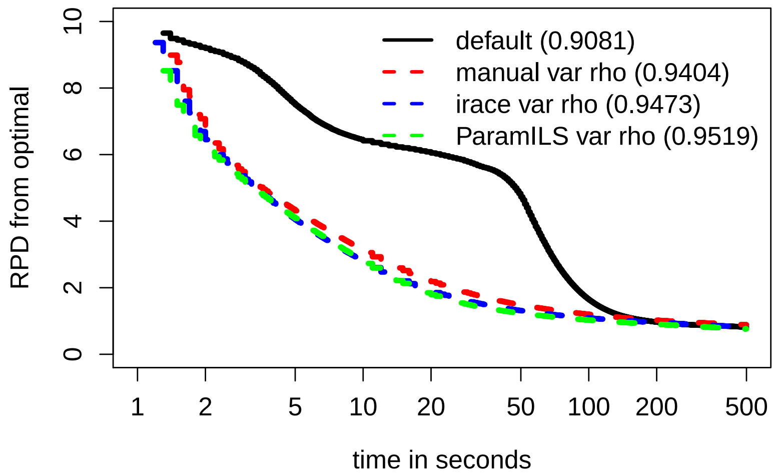
Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043



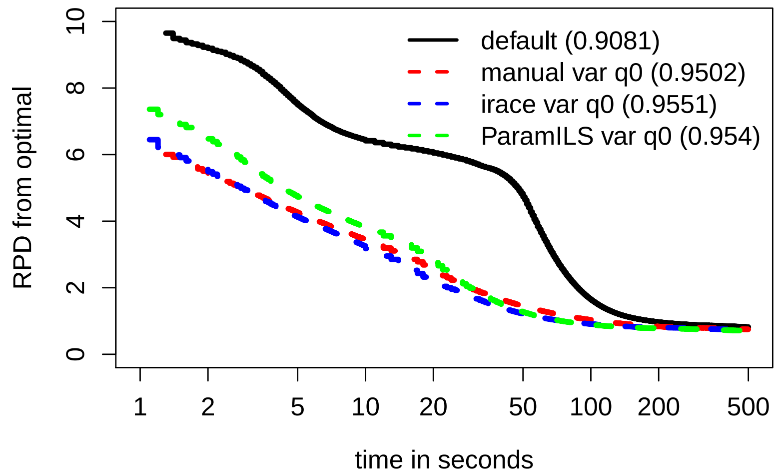
(a) Parameter m (ants)



(b) Parameter β (beta)



(c) Parameter ρ (rho)



(d) Parameter q_0

Figure 5: Anytime behaviour of manually tuned vs. automatically tuned configurations of MMAS. The two automatically tuned configurations were obtained by using either `irace` or ParamILS, both with the same tuning budget (1 000 runs of MMAS). The number in the legend is the mean hypervolume of each configuration over all runs. The plots show that both `irace` and ParamILS obtain configurations with very similar anytime behaviour and both are able to surpass both the default and the manually tuned configurations.

Please cite this article as: López-Ibáñez, M., & Stützle, T. Automatically improving the anytime behaviour of optimisation algorithms. European Journal of Operational Research, 235(3): 569-582 (2014), doi: 10.1016/j.ejor.2013.10.043

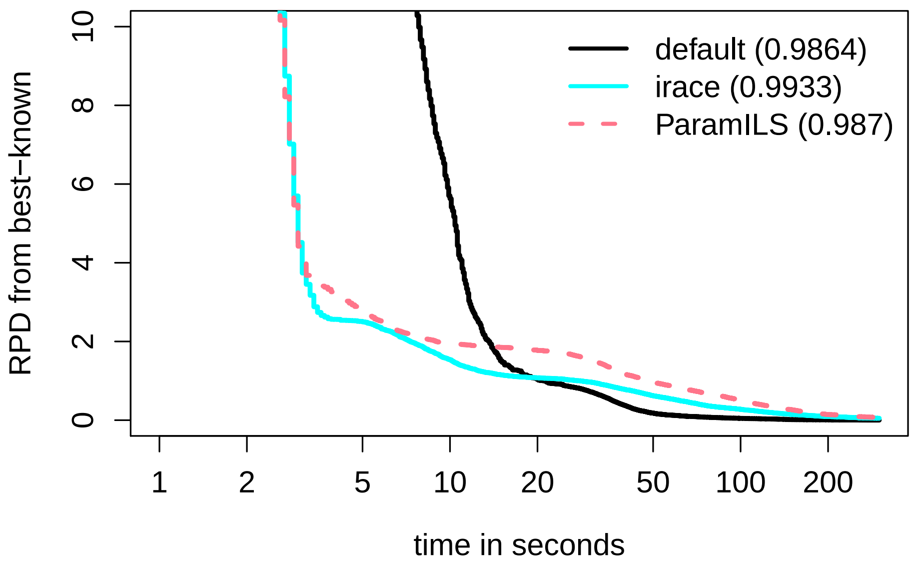


Figure 6: Mean RPD over all test instances for different configurations of SCIP: the default configuration and two configurations automatically tuned for anytime behaviour using the classical hypervolume by means of either `irace` or ParamILS. The number in parentheses is the mean hypervolume corresponding to that configuration. In this case, the anytime behaviour of the configuration obtained by ParamILS is worse than the one obtained by `irace`.