Vol 1, No 1, pp. 38–40, 2009
International Journal of
Mathematical and Computational
Forestry & Natural-Resource Sciences

AVAILABLE ONLINE AT HTTP://MCFNS.COM

Submitted: Dec. 12, 2008
Accepted: Dec. 24, 2008
Published: Feb. 28, 2009
Last Correction: Mar. 09, 2009

A NOTE ON "A REVIEW OF THE STATUS AND USE OF VALIDATION PROCEDURES FOR HEURISTICS USED IN FOREST PLANNING"

Alan T. Murray

Professor, SGS, Arizona State University, Tempe AZ 85287 USA. Ph./FAX: 480.965.7533/8313

Abstract. This paper provides a discussion on validation procedures associated with heuristic solution approaches used in forest planning, initiated by Bettinger et al. (2008) (Bettinger, Sessions and Boston, 2008. *A review of the status and use of validation procedures for heuristics used in forest planning.* MCFNS 1(1): 26–37). Three issues are addressed.

Keywords: Heuristics, forest planning, validation

1 Background

Bettinger et al. (2008) promotes a public dialogue about the academic significance of developed heuristics in solving forest planning problems. As suggested in their paper, the development, use and evolution of heuristics to solve forestry problems has been significant over the past few decades. It is therefore timely to reflect on this field of research. Their paper suggests that there are different cases for considering/evaluating developed heuristics, making potential research worthy of publication in many different possible contexts. This is noteworthy and important as there is sometimes a failure to understand and appreciate the nuances of a particular heuristic application. I have seen this as an author, referee and editor, and have also heard from other journal editors over the years that research involving advances in heuristic solution techniques is difficult to assess. In fact, I recall an editor suggesting in the late 1990s that formal guidance for assessment in forestry applications could be helpful.

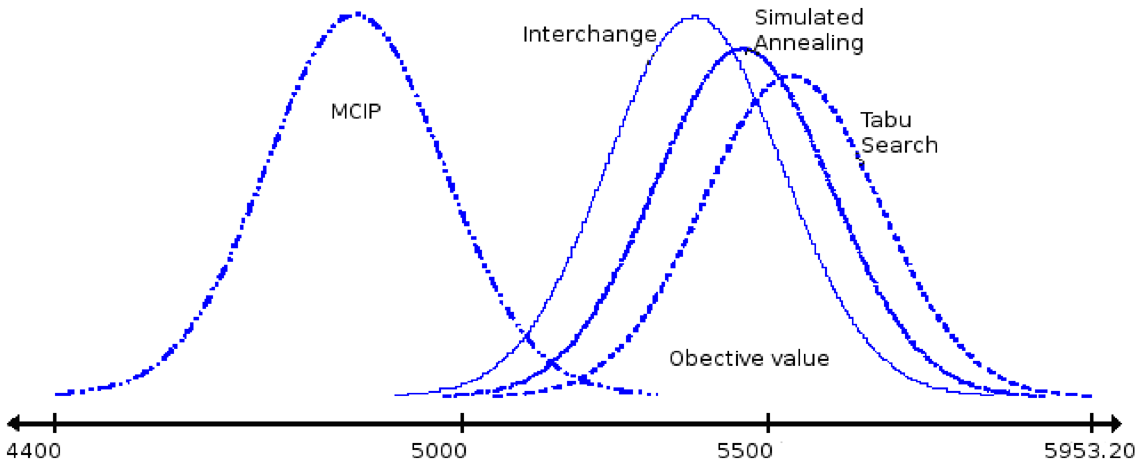
I have been invited to provide this discussion because of three observations regarding Bettinger et al. (2008). First, the issue concerning heuristic research as an important sub-area has received much attention in the academic community outside of forestry. Second, statistical significance is far from established in the application of heuristics to solve optimization problems. Third, there are potential problems with suggesting "levels of validation". I elaborate on these points below.

2 Discussion

The first issue that comes across in the paper is an argument that heuristic development to solve forest planning problems is a valid and publication worthy subfield. I believe that it is. In fact, it is recognized as such in the broader academic community, especially in the area of operations research. To this end, the paper is not new or novel in what is being said, but perhaps this is something that the forestry community needs to hear. The area of heuristic application and development has always had to contend with the fact that one cannot attribute any statistical or qualitative indicator to a heuristically generated solution(s) as there is generally no basis for establishing certainty. While this is true, heuristics continue to be needed to solve difficult planning problems and may in fact be more prominent in the future.

The second issue relates to the statistical significance as discussed with "Level 4" validation. In my opinion the paper comes across as overly optimistic about establishing an "estimated global optimum solution." Heuristics are necessarily biased in sampling solution space because they generally attempt to identify solutions that are toward the tail of the associated feasible solution distribution, or rather are "close" to optimal. Another issue that thwarts the possibility of statistical significance is independence in a sample of solutions. The literature on heuristics for optimization problems is dotted with examples where researchers have observed similarity or overlap in identified solutions. In fact, the more recent emergence of the so-called heuristic concentration ap-

Figure 1: Comparative characteristics of feasible solutions identified using alternative heuristic approaches.



proach of Rosing and ReVelle (1997) is a prime example. Heuristic concentration works by first identifying a number of "good" solutions using an approximate (heuristic) technique, then searches from among the identified decision variables equal to one in value. What is found is that it is possible to find optimal and near optimal solutions using the approach. A concern related to statistical significance is validity in the context of optimization. My own research in the use, development and application of heuristics for solving forest planning problems is that it is virtually impossible to infer statistical significance in a general sense. As an example, the comparative study reported in Murray and Church (1995) developed three new heuristic approaches (interchange, simulated annealing and tabu search) to solve a forest planning problem with adjacency, green up, temporal volume and roading considerations. Results were compared with the monte carlo integer programming (MCIP) heuristic approach of Nelson and Brodie (1990) as well as the optimum derived using mixed integer programming. A simplified summary of the 1300 derived solutions[1] for each heuristic reported in Murray and Church (1995) is shown in Figure 1. The horizontal axis represents the objective value of an identified solution, and the vertical axis is the relative frequency of an observed value in terms of the sample of solutions. Is it possible to infer anything of value from the observed distributions in any of the four cases? It is clear that there are differences in the performance of the different heuristics, and higher observed values are better, but does that establish any certainty or statistical significance about the quality of a given solution? Absolutely not in these individual cases, and this is precisely the issue that heuristic solution techniques are faced with. As an example, if one only applied the MCIP approach and had no other knowledge, the best solution found might appear to be of very high quality and statistically significant. However, compared to the samples identified using the other heuristics shown in Figure 1 (interchange, simulated annealing and tabu search), this would not be so sensible. If I disclose that the optimal solution is 5953.20 (indicated on the horizontal axis), then qualitative assessment is possible but not beyond this particular problem application. Thus, concluding something about solution quality for a tabu search solution identified for a harvest scheduling problem, as an example, is not possible and any attempt to do so would be invalid and misleading relative to the true optimal solution. It is not possible to infer any kind of statistical significance because the actual distribution of feasible (or even infeasible) solutions is not generally known. Further, good heuristics necessarily sample the tail of the distribution anyway, so this complicates matters even more.

Finally, the third issue is that the "levels of validation" reflect an attempt to classify approaches that have been used in the literature, but some researchers may be tempted to use them to justify a particular application or developed heuristic. It really suggests (or concludes) that if your work falls into one or more of the cases, then achieving the stipulated levels means it is publishable. Unfortunately the reality of research is that things are never so cut and dry. After all, what is a "new" problem? Is it a "standard" problem with a slight twist, say one additional constraint of some kind? What about a "new" heuristic? Does a "standard" heuristic with a

[1] The harvest scheduling problem was to maximize economic return, having 291 binary decision variable (45 management units, 52 road segments and three 3 periods). Thus, the number of potential solutions is 2^{291}, though many are no doubt infeasible.

minor or major change/interpretation qualify as new? This is not at all clear, and necessarily requires subjective evaluation.

3 Conclusion

While I support the intent and overview provided by Bettinger et al. (2008), I do feel the above comments are important issues to keep in mind associated with the development, use and application of heuristic methods to solve forest planning problems . In the spirit of academic dialogue and the fact that evaluating heuristic applications in forest planning is challenging in the publication process, the review that they provide is potentially useful and valuable for researchers to consider. Having said this, the fact remains that there are well established approaches for “validating” heuristic based research, as noted in Bettinger et al. (2008) as well as included citations. The paper does a reasonable job describing potential approaches that have been utilized in the literature. Nevertheless, the assessment of research (peer review) relies on individuals that may or may not be qualified, or inclined, to review work, and is necessarily subjective.

Acknowledgement

This note has benefited from comments provided by referees and the editor.

References

Bettinger, P., J. Sessions, K. Boston. 2008. A review of the status and use of validation procedures for heuristics used in forest planning. International Journal of Mathematical and Computational Forestry & Natural-Resourece Sciences 1: 21–32.

Murray, A.T., R.L. Church. 1995. Heuristic solution approaches to operational forest planning problems. OR Spektrum 17: 193–203.

Nelson, J., J.D. Brodie. 1990. Comparison of a random search algorithm and mixed integer programming for solving area-based forest plans. Canadian Journal of Forest Research 20: 934–942.

Rosing, K.E., C.S. ReVelle. 1997. Heuristic concentration: two stage solution construction. European Journal of Operational Research 97: 75–86.