**Title:** Marxan vs ILP

**Authors:** Schuster R.a,b,\*, Matt Strimas-Mackeyc,, Jeffrey O. Hansond, Bennett, J. Ra

a Department of Biology, 1125 Colonel By Drive, Carleton University, Ottawa ON, K1S 5B6 Canada.

b Ecosystem Science and Management Program, 3333 University Way, University of Northern British Columbia, Prince George BC, V2N 4Z9 Canada.

c Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850 USA.

d School of Biological Sciences, The University of Queensland, Brisbane, QLD 4072, Australia

**\*** Corresponding author:Richard Schuster, Department of Biology, 1125 Colonel By Drive, Carleton University, Ottawa ON, K1S 5B6 Canada. 250-635-2321. richard.schuster@glel.carleton.ca

**Keywords**: Marxan, Integer Linear Programming, Optimization, Prioritization

**Abstract**

**Introduction**

Systematic conservation planning (SCP) is a rigorous, repeatable, and structured approach to designing new protected areas that efficiently meet conservation objectives (Margules and Pressey 2000). Historically, conservation decision-making has often evaluated parcels opportunistically as they became available for purchase, donation, or under threat. Although purchasing such areas may improve the status quo, such decisions may not substantially enhance the long-term persistence of target species or communities. SCP is a systematic alternative to this opportunistic approach, using decision support tools to simulate alternative reserve designs over a range of different biodiversity and management goals and, ultimately, guide protected area acquisitions and management actions. Due to the systematic, evidence-based nature of these tools, conservation prioritization can help contribute to a transparent, inclusive, and more defensible decision-making process.

There are two main approaches to solving optimization problems of this type. First, integer linear programming (ILP), which minimizes or maximizes an objective function (a mathematical equation describing the relationship between actions and out-comes) subject to a set of constraints and conditional on the decision variables (the variables corresponding to the selection of actions to implement) being integers. Second, solutions can be found using heuristic methods such as simulated annealing (SA; Kirkpatrick et al., 1983), which iteratively, stochastically explore the state-space of the decision variables. There are numerous other heuristics (e.g. ranking procedures, genetic algorithms, and mixtures of these approaches) that could also be used. Here, we focus on SA because it is the most widely used heuristic in the conservation planning literature in the form of the conservation planning software Marxan (Ball et al., 2009; Watts et al., 2009) and, unlike deterministic heuristics such as ranking, it is possible that SA could find an optimal solution to any problem.

Marxan is the most widely used SCP software globally, being used in 184 countries to design marine and terrestrial reserve systems and is the global leader in conservation land and sea use planning software. Marxan uses simulated annealing, a heuristic optimization technique, to find ‘near optimal’ solutions to SCSP problems. Recent developments in computational capacity and algorithms has made it possible to solve the SCP problems Marxan solves with integer linear programming (Beyer et al. 2016). Building on Beyer et al. (Beyer et al. 2016), we created a software package for the R statistical software called prioritizr, that can solve Marxan type problems using integer linear programming (Hanson et al. 2019).

Here, we are using a case study from Western North America to compare Marxan (simulated annealing) and prioritizr (integer linear programming) to ask the following questions:

1. What is required to parameterize i) Marxan, ii) prioritizr using an open source solver, and iii) prioritzr using a proprietary solver?
2. How do processing time differ between the three approaches tested?
3. How cost effective, in $ values, are the three approaches tested?

**Methods**

*Study area*

We focused on a 27,250 km2 portion of the Georgia Basin, Puget Trough and Willamette Valley of the Pacific Northwest region spanning the US and Canada (Fig. 1), corresponding to the climate envelope indicative of the Coastal Douglas-fir (CDF) Biogeoclimatic zone in southwestern British Columbia (Meidinger and Pojar 1991). Land cover in the region is diverse, with approximately 57% of the land in forest, 8% as savanna or grassland, 5% in cropland, and 10% being urban or built.

*Data Layers*

*Biodiversity data. We used species distribution models for 72 bird species as our conservation features.* The distribution models were based on data from eBird, a citizen-science effort that has produced the largest and most rapidly growing biodiversity database in the world (Hochachka et al. 2012, Sullivan et al. 2014). From the 2013 eBird Reference Dataset (<http://ebird.org/ebird/data/download>) we used a total of 12081 checklists in our study area, then filtered these checklists to retain only those <1.5 hours in duration, <5 km travelled, and a maximum of 10 visits to a given location (unpublished R code; Hochachka, pers. com.). Sampling locations <100 m apart were collapsed to one location, yielding 5470 checklists from 2160 locations, visited from 1-10 times and 2.53 times on average. unmarked. Following (Schuster et al. 2014, 2017) we used a combination of quantitative models and expert elicitation to identify which species were associated either with forest habitat or with human-dominated habitat, such as built or residential land (Supplementary Table 1). For further details see Rodewald et al. (XXXX).

*Cadastral layer and land cost*. We incorporated spatial heterogeneity in land cost (Ando et al. 1998, Polasky et al. 2001, Ferraro 2003, Naidoo et al. 2006) in our plan by using cadastral data and 2012 land value assessments from the Integrated Cadastral Information Society of BC, resulting in 193,623 polygons for BC (Schuster et al. 2014). Cadastral data, including tax assessment land values from Washington State came from the University of Washington’s Washington State Parcel Database (<https://depts.washington.edu/wagis/projects/parcels/>; Version: StatewideParcels\_v2012n\_e9.2\_r1.3; Date accessed: 2015/04/30), as well as San Juan County Parcel Data with separate signed user agreement. The combined cadastral layer included 1.92M polygons. Cadastral data, including tax assessment land values from Oregon State had to be sourced from individual counties, which included Benton, Clackamas, Columbia, Douglas, Lane, Linn, Marion, Multnomah, Polk, Washington and Yamhill. The combined cadastral layer for Oregon included 605,425 polygons.

*Spatial prioritization approach*

Here we use the concept of systematic conservation planning (Margules and Pressey 2000), to inform choices about areas to protect, in order to optimize outcomes for biodiversity while minimizing societal costs (McIntosh et al. 2017). To achieve the goal to optimize the trade-off between conservation benefit and socioeconomic cost, i.e. to get the most benefit for limited conservation funds, we strive to minimize an objective function over a set of decision variables, subject to a series of constraints.

Marxan formulation

Integer linear programming (ILP) is the subset of optimization algorithms used here to solve reserve design problems. The general form of an ILP problem can be expressed in matrix notation as:

Where x is a vector of decision variables (in our case, whether to prioritize an individual planning unit), c and b are vectors of known coefficients, and A is the constraint matrix. In the minimum set cover problem, c is a vector of costs for each planning unit, b a vector of targets for each conservation feature, the relational operator would be ≥ for all features, and A is the representation matrix with Aij=rij, the representation level of feature i in planning unit j. We set an objective to find the solution that fulfills all the targets and constraints for the smallest area, which we use as our measure of cost (Beyer et al. 2016).

*Scenarios investigated*

We investigated a range of scenarios that were computationally feasible for this study. For both Marxan and prioritzr scenarios we created the following range of scenarios: i) vary conservation targets between 10 and 90 % in 10 % increments (9 variations), ii) 10 – 72 species/features (5 variations) and iii) 9282, 37128, 148510 planning units (3 variations), resulting in a total of 135 scenarios created. For Marxan we also varied two additional parameters, i) number of iterations from 1E+04 to 1E+08 (5 iterations) and ii) the species penalty factor 1, 5, 25, 125 (4 variations) for a total of 2700 scenarios investigated in Marxan. As the processing time for the most complex problem in Marxan (90% target, 72 features, 148510 planning units, 1E+08 iterations) already took 8 hours to solve, we restricted the set of full scenario iterations to those mentioned above. We did however go ahead and create an additional 9 scenarios (target range from 10 – 90 %, with 72 features, 1E+08 iterations and spf = 5) to present as one of the main results for comparison here, as the next higher iterations of number of planning units (n = 594040) does best highlight the limitations of Marxan/simulated annealing compared to prioritzr/integer linear programming, when it comes to finding the optimal solution and being most cost effective in selecting planning units.

**Results**

The best processing time were achieved using the prioritizr package and the commercial solver Gurobi, followed by prioritizr and the open source solver Symphony, and lastly Marxan (Figure 1). Gurobi was as fast or faster across all scenarios investigated, Symphony took between 0 and 113 times longer than Gurobi (mean = 18.4 times), Marxan took between 0 and 28710 times longer than Gurobi (mean = 1071 times).

Integer linear programming algorithms (Gurobi, Symphony) outperformed simulated annealing (Marxan) in terms of finding the optimal solution in every single case. This resulted in a lower objective value, but in our case of using assessed land values as cost, we show that cost savings ranging from 12 to 30% result in hugely reduced expenditures. At the 30% protection target ILP solvers resulted in solutions that were $144M cheaper than SA.

**Discussion**

**Conclusion**

Integer linear programming algorithms outperform simulated annealing as used in Marxan substantially, both in terms of time required to find near optimal or optimal solutions and more importantly in terms of solution cost. Using an integer linear programming algorithm, as implemented in the R package prioritizr, has the added benefit that users don’t need to worry or set parameters such as species penalty factors or number of iterations anymore, which significantly the time a user spends of finding suitable values for these parameters. With the capabilities of prioritizr, including everything Marxan can do and more, we highly recommend users adopting this modified approach to solving systematic conservation planning problems.

**Acknowledgements**

RS is supported by a Liber Ero Fellowship, MSM by endowments at the Cornell Lab of Ornithology, JOH by XXX , and JRB by Natural Sciences and Engineering Research Council of Canada and Environment and Climate Change Canada.

**References**

Ando, A., J. Camm, S. Polasky, and A. Solow. 1998. Species Distributions, Land Values, and Efficient Conservation. Science 279:2126–2128.

Ball, I. R. R., H. P. P. Possingham, and M. E. E. Watts. 2009. Marxan and relatives: Software for spatial conservation prioritisation. Pages 185–195 *in* A. Moilanen, K. Wilson, and H. P. Possingham, editors. Spatial conservation prioritisation: Quantitative methods and computational tools. Oxford University Press, Oxford.

Beyer, H. L., Y. Dujardin, M. E. Watts, and H. P. Possingham. 2016. Solving conservation planning problems with integer linear programming. Ecological Modelling 328:14–22.

Ferraro, P. J. 2003. Assigning priority to environmental policy interventions in a heterogeneous world. Journal of Policy Analysis and Management 22:27–43.

Hanson, J., R. Schuster, N. Morrell, M. Strimas-Mackey, M. E. Watts, P. Arcese, J. R. Bennett, and H. P. Possingham. 2019. prioritizr: Systematic Conservation Prioritization in R, Version 4.0.2.

Hochachka, W. M., D. Fink, R. A. Hutchinson, D. Sheldon, W.-K. Wong, and S. Kelling. 2012. Data-intensive science applied to broad-scale citizen science. Trends in ecology & evolution 27:130–137.

Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. Nature 405:243–53.

McIntosh, E. J., R. L. Pressey, S. Lloyd, R. Smith, and R. Grenyer. 2017. The Impact of Systematic Conservation Planning. Annual Review of Environment and Resources 42:annurev-environ-102016-060902.

Meidinger, D., and J. Pojar. 1991. Ecosystems of British Columbia. British Columbia Ministry of Forests, Victoria, BC.

Naidoo, R., A. Balmford, P. J. Ferraro, S. Polasky, T. H. Ricketts, and M. Rouget. 2006. Integrating economic costs into conservation planning. Trends in ecology & evolution 21:681–7.

Polasky, S., J. D. Camm, and B. Garber-Yonts. 2001. Selecting Biological Reserves Cost-Effectively: An Application to Terrestrial Vertebrate Conservation in Oregon. Land Economics 77:68–78.

Schuster, R., E. A. Law, A. D. Rodewald, T. G. Martin, K. A. Wilson, M. Watts, H. P. Possingham, and P. Arcese. 2017. Tax Shifting and Incentives for Biodiversity Conservation on Private Lands. Conservation Letters.

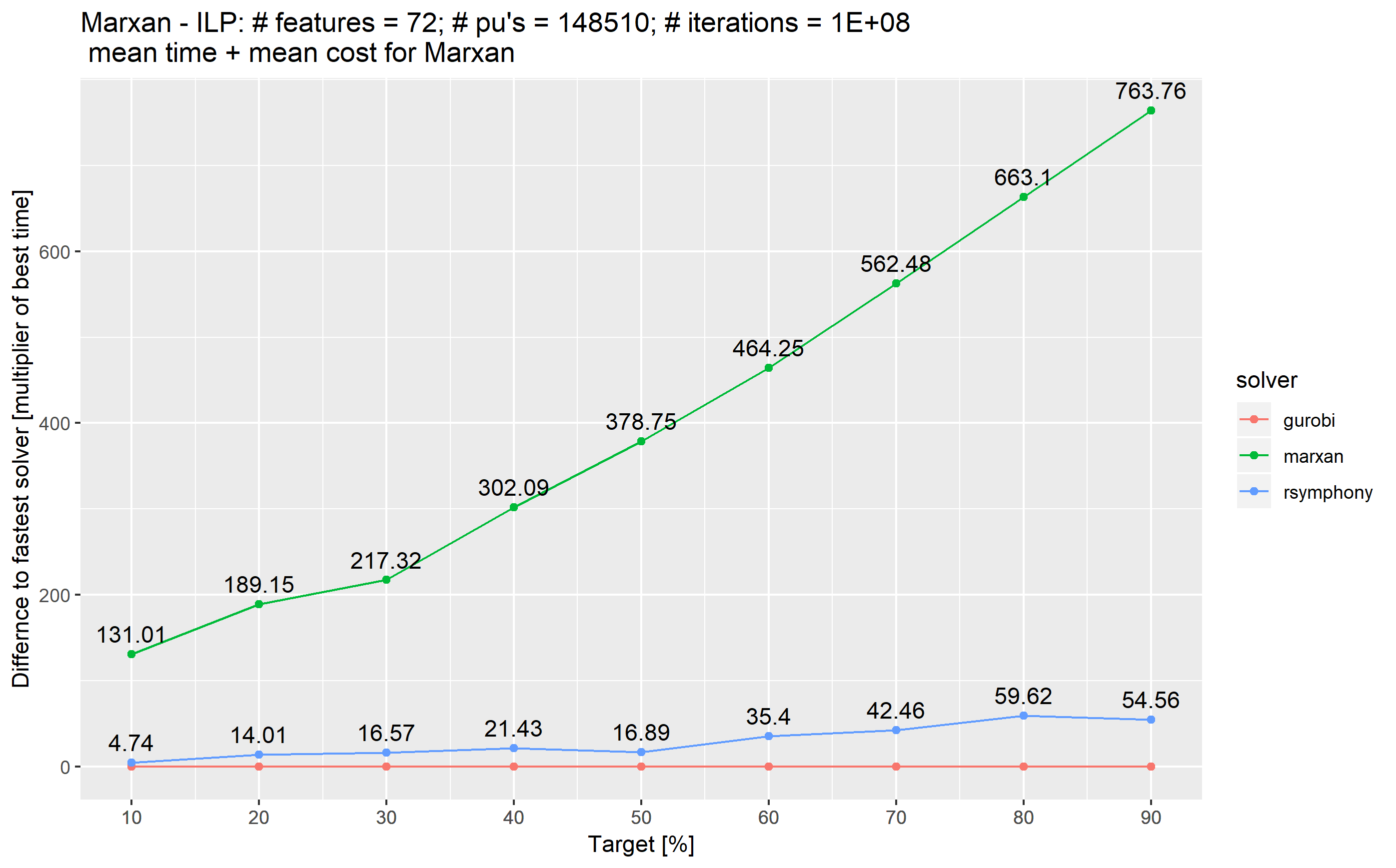
Schuster, R., T. G. Martin, and P. Arcese. 2014. Bird community conservation and carbon offsets in Western North America. PLoS ONE.

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, and others. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. Biological Conservation 169:31–40.

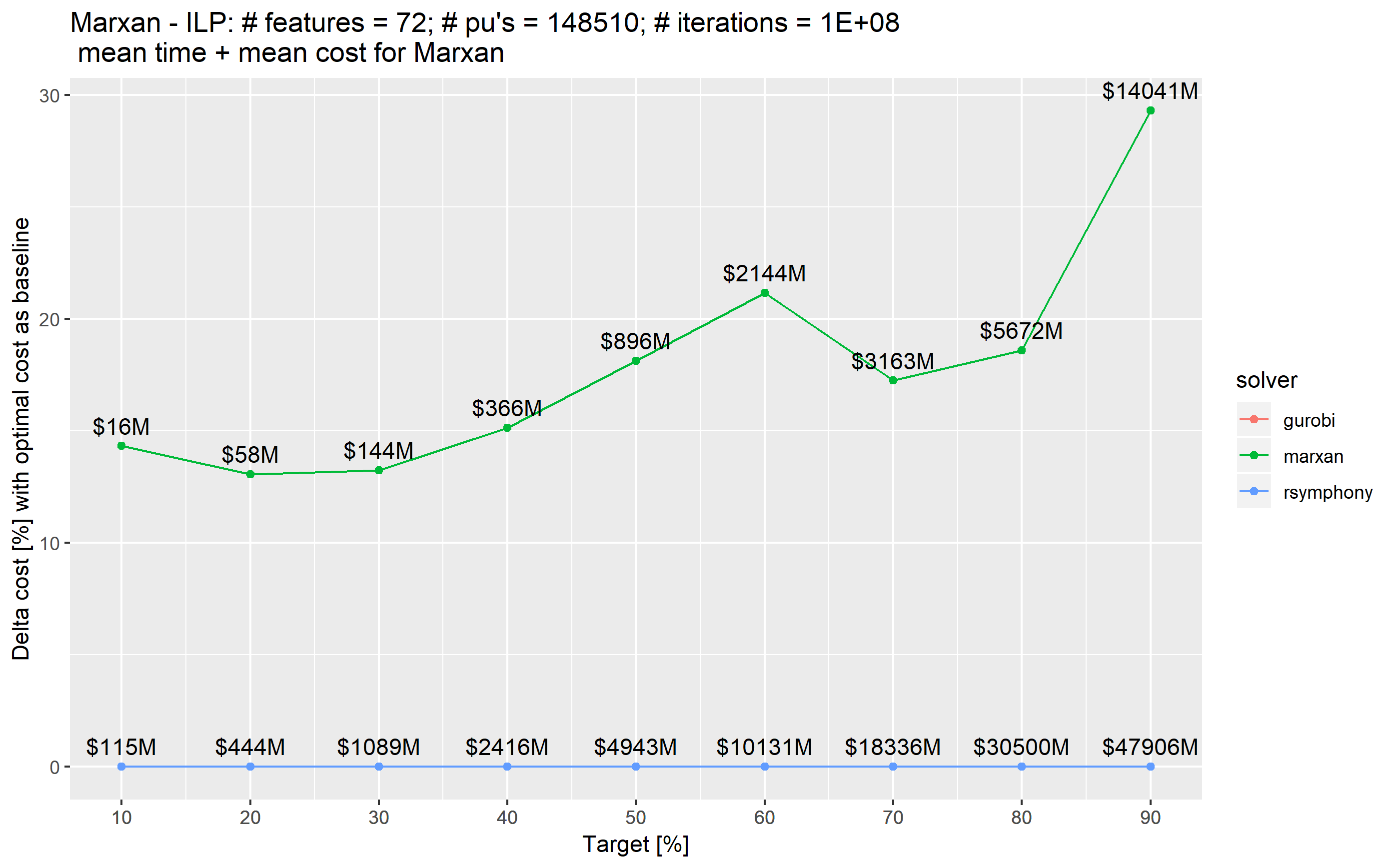
**Table 1.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Paremeter** | **Value range** | **n** | **Scenarios** |
| targets | 10 - 90% | 9 |  |
| n features | 10, 26, 41, 56, 72 | 5 |  |
| n pu | 9282, 37128, 148510 | 3 | 135 (ILP) |
| marxan iterations | 1E+04, 1E+05, 1E+06, 1E+07, 1E+08 | 5 |  |
| marxan spf | 1, 5, 25, 125 | 4 | 2700 (Marxan) |

**Figure 1.**

****

**Figure 2.**

****

**Supplementary Table 1.**

|  |  |  |
| --- | --- | --- |
| Species Code | Common Name | Scientific Name |
| amegfi | American Goldfinch | Spinus tristis |
| amekes | American Kestrel | Falco sparverius |
| amerob | American Robin | Turdus migratorius |
| annhum | Anna's Hummingbird | Calypte anna |
| baleag | Bald Eagle | Haliaeetus leucocephalus |
| barswa | Barn Swallow | Hirundo rustica |
| brdowl | Barred Owl | Strix varia |
| belkin1 | Belted Kingfisher | Megaceryle alcyon |
| bewwre | Bewick's Wren | Thryomanes bewickii |
| bnhcow | Brown-headed Cowbird | Molothrus ater |
| bkhgro | Black-headed Grosbeak | Pheucticus melanocephalus |
| brebla | Brewer's Blackbird | Euphagus cyanocephalus |
| brncre | Brown Creeper | Certhia americana |
| batpig1 | Band-tailed Pigeon | Patagioenas fasciata |
| bushti | Bushtit | Psaltriparus minimus |
| cangoo | Canada Goose | Branta canadensis |
| chbchi | Chestnut-backed Chickadee | Poecile rufescens |
| cedwax | Cedar Waxwing | Bombycilla cedrorum |
| chispa | Chipping Sparrow | Spizella passerina |
| coohaw | Cooper's Hawk | Accipiter cooperii |
| comrav | Common Raven | Corvus corax |
| amecro | American Crow | Corvus brachyrhynchos |
| dowwoo | Downy Woodpecker | Dryobates pubescens |
| eucdov | Eurasian Collared-Dove | Streptopelia decaocto |
| eursta | European Starling | Sturnus vulgaris |
| evegro | Evening Grosbeak | Coccothraustes vespertinus |
| norfli | Northern Flicker | Colaptes auratus |
| foxspa | Fox Sparrow | Passerella iliaca |
| gockin | Golden-crowned Kinglet | Regulus satrapa |
| haiwoo | Hairy Woodpecker | Dryobates villosus |
| houfin | House Finch | Haemorhous mexicanus |
| houspa | House Sparrow | Passer domesticus |
| houwre | House Wren | Troglodytes aedon |
| hutvir | Hutton's Vireo | Vireo huttoni |
| macwar | MacGillivray's Warbler | Geothlypis tolmiei |
| moudov | Mourning Dove | Zenaida macroura |
| norhar1 | Hen Harrier | Circus cyaneus |
| orcwar | Orange-crowned Warbler | Oreothlypis celata |
| olsfly | Olive-sided Flycatcher | Contopus cooperi |
| osprey | Osprey | Pandion haliaetus |
| pacwre1 | Pacific Wren | Troglodytes pacificus |
| pinsis | Pine Siskin | Spinus pinus |
| pilwoo | Pileated Woodpecker | Dryocopus pileatus |
| pasfly | Pacific-slope Flycatcher | Empidonax difficilis |
| purfin | Purple Finch | Haemorhous purpureus |
| purmar | Purple Martin | Progne subis |
| rebnut | Red-breasted Nuthatch | Sitta canadensis |
| rebsap | Red-breasted Sapsucker | Sphyrapicus ruber |
| redcro | Red Crossbill | Loxia curvirostra |
| rocpig | Rock Pigeon | Columba livia |
| rethaw | Red-tailed Hawk | Buteo jamaicensis |
| rufhum | Rufous Hummingbird | Selasphorus rufus |
| rewbla | Red-winged Blackbird | Agelaius phoeniceus |
| savspa | Savannah Sparrow | Passerculus sandwichensis |
| sora | Sora | Porzana carolina |
| sonspa | Song Sparrow | Melospiza melodia |
| spotow | Spotted Towhee | Pipilo maculatus |
| stejay | Steller's Jay | Cyanocitta stelleri |
| swathr | Swainson's Thrush | Catharus ustulatus |
| towwar | Townsend's Warbler | Setophaga townsendi |
| treswa | Tree Swallow | Tachycineta bicolor |
| daejun | Dark-eyed Junco | Junco hyemalis |
| yerwar | Yellow-rumped Warbler | Setophaga coronata |
| varthr | Varied Thrush | Ixoreus naevius |
| vigswa | Violet-green Swallow | Tachycineta thalassina |
| warvir | Warbling Vireo | Vireo gilvus |
| whcspa | White-crowned Sparrow | Zonotrichia leucophrys |
| westan | Western Tanager | Piranga ludoviciana |
| wilsni1 | Wilson's Snipe | Gallinago delicata |
| wlswar | Wilson's Warbler | Cardellina pusilla |
| wooduc | Wood Duck | Aix sponsa |
| yelwar | Yellow Warbler | Setophaga petechia |