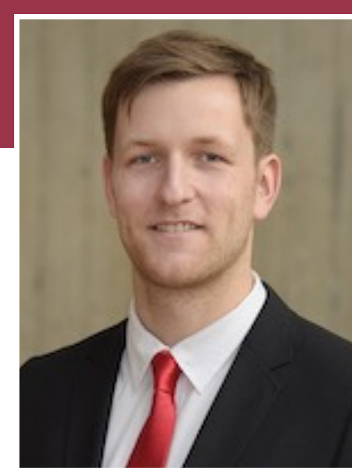
Automated and Feature-Based Algorithm Selection on Single-Objective Continuous Black-Box Problems



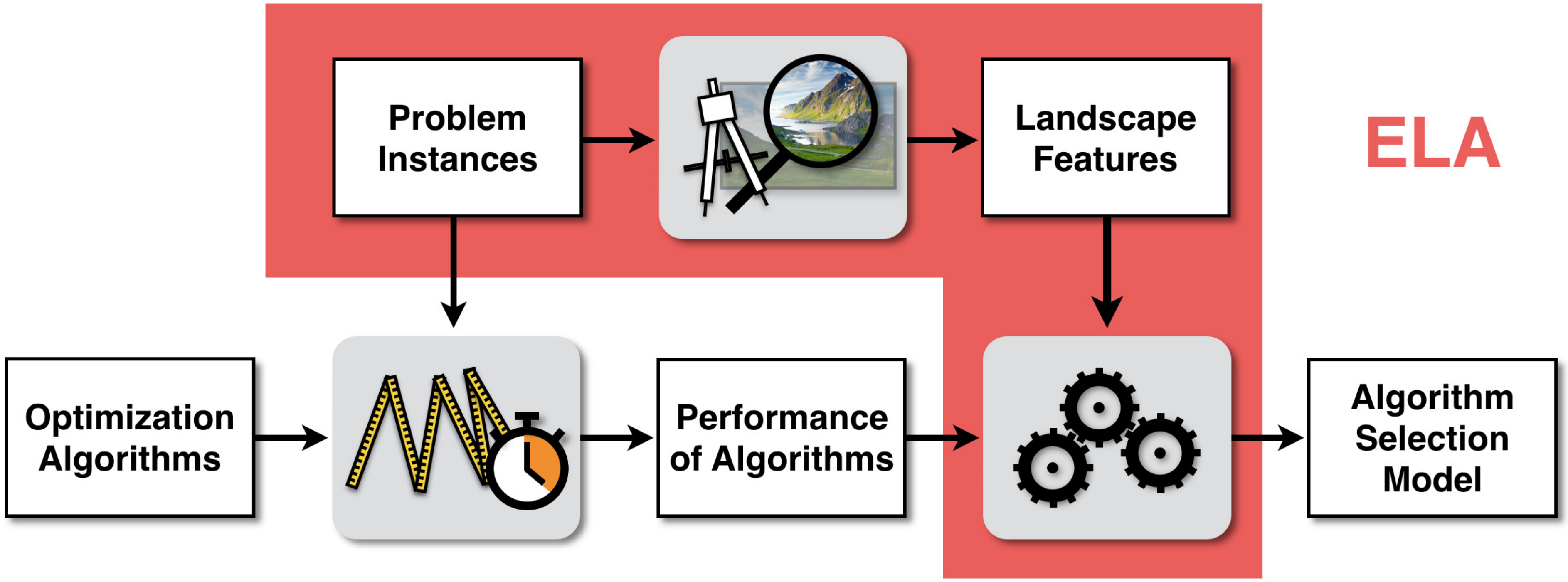
Pascal Kerschke
University of Münster

Heike Trautmann
University of Münster

Presentation at the 5th COSEAL-Workshop
Monday, September 11, 2017
in Brussels, Belgium

Context:

- **Algorithm Selection**
  John R. Rice, „The Algorithm Selection Problem," Advances in Computers, vol. 15, pp. 65 – 118, 1976.
- **Continuous Black-Box Optimization (Benchmark)**
  Nikolaus Hansen, Anne Auger, Steffen Finck, and Raymond, Ros, „Real-Parameter Black-Box Optimization Benchmarking 2009: Experimental Setup," INRIA, Tech. Rep. RR-6828, 2009.
- **Problem Dependent Landscape-Features**



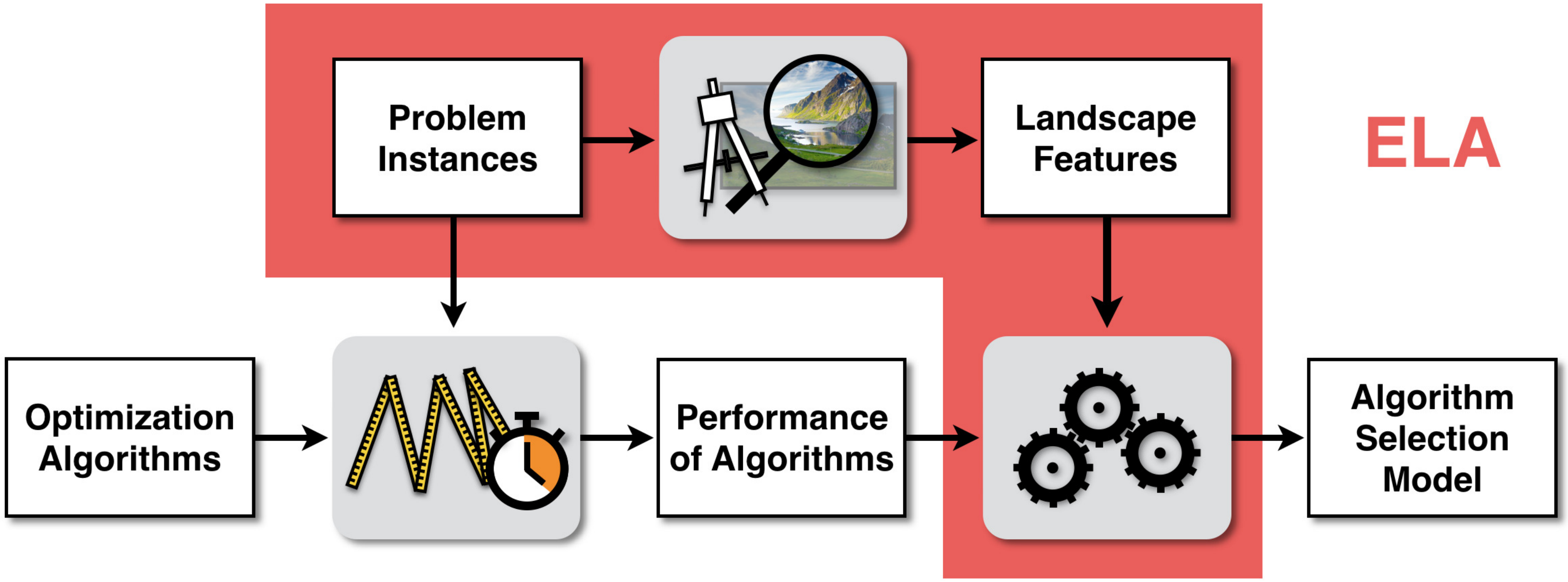
Exploratory Landscape Analysis (ELA):

- **initial work:**

  Olaf Mersmann, Bernd Bischl, Heike Trautmann, Mike Preuss, Claus Weihs, and Günter Rudolph, „Exploratory Landscape Analysis," in Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, 2011, pp. 829 – 836.

- **sophisticated and *automated* approach for characterizing landscapes**

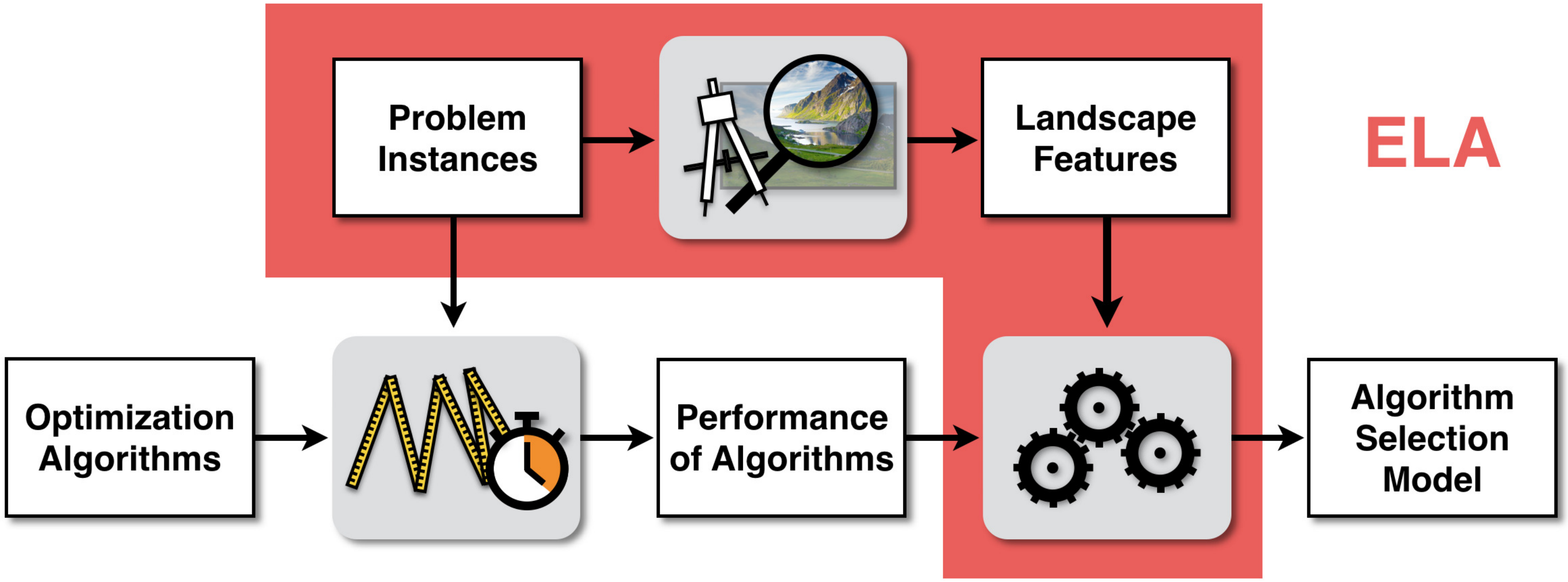
- **single-objective continuous optimization problems**



Exploratory Landscape Analysis (ELA):

- different feature sets can be useful for different landscape properties
- low-budget ELA: able to compute landscape features with a budget of 50 x d function evalutions (d = problem dimensionality)

Pascal Kerschke, Simon Wessing, Mike Preuss, Heike Trautmann, „Low-Budget Exploratory Landscape Analysis on Multiple Peaks Models," in Proceedings of the 18th Annual Conference on Genetic and Evolutionary Computation (GECCO). ACM, 2016, pp. 229 – 236.



Experimental Data

- **FLACCO: An R-Package for ELA**
- **Feature-Based Landscape Analysis of Continuous and Constrained Optimization Problems**

  Pascal Kerschke. *Comprehensive Feature-Based Landscape Analysis of Continuous and Constrained Optimization Problems Using the R-Package flacco*. In: Journal of Statistical Software (under review). Current version of the manuscript is available at arXiv: https://arxiv.org/abs/1708.05258

- **collection of 17 different feature sets (i.e., more than 300 features)**
- **tracks costs (# function evaluations and runtime) per feature set**
- **additionally provides multiple visualization techniques**
- **package releases and further online material:**
  - stable release (vs. 1.7) on CRAN: https://cran.r-project.org/package=flacco
  - developers version (vs. 1.7) on GitHub: https://github.com/kerschke/flacco
  - online tutorial: http://kerschke.github.io/flacco-tutorial/site/
  - GUI: https://flacco.shinyapps.io/flacco/

Experimental Data

- considered problem dimensions: $d \in \{2, 3, 5, 10\}$
- computed 102 features per problem instance:
  - „classical“ ELA features (convexity, curvature, levelset, local search, meta-model and y-distribution)
  - cell mapping angle features
  - dispersion features
  - information content features
  - nearest better clustering features
  - basic features
  - principal component features
- used initial design of 50 x d observations (i.e., 100 to 500)

Experimental Data

- used performance data from the COCO-platform

  Nikolaus Hansen, Anne Auger, Olaf Mersmann, Tea Tušar, and Dimo Brockhoff. *COCO: A Platform for Comparing Continuous Optimizers in a Black-Box Setting*. Available at arXiv: http://arxiv.org/abs/1603.08785v3

- collection of performance results from 129 optimization algorithms
- performance measured by *Expected Runtime* (and PAR10)
- COCO provides pairs of function evaluations and reached objective value
- community usually considers fixed ***absolute*** precision threshold of 10^{-5} to 10^{-7}
- here: precision threshold of 10^{-2}

  (approx. 67% of all runs terminated successfully)

Experimental Data

- algorithms executed on BBOB problems

  in context of BBOB competitions, hosted at GECCO in the years 2009, 2010, 2012, 2013 and 2015

- considered instances (IIDs) changed per competition year:

  2009: 1 to 5 (3 replications)
  2010: 1 to 15 (1 replication)
  2012: 1 to 5 and 21 to 30 (1 replication)
  2013: 1 to 5 and 31 to 40 (1 replication)
  2015: 1 to 5 and 41 to 50 (1 replication)

- forced to use IIDs 1 to 5 (1 replication)

- compute ERT on FID level $\Rightarrow$ 96 problems (4 dimensions à 24 FIDs)

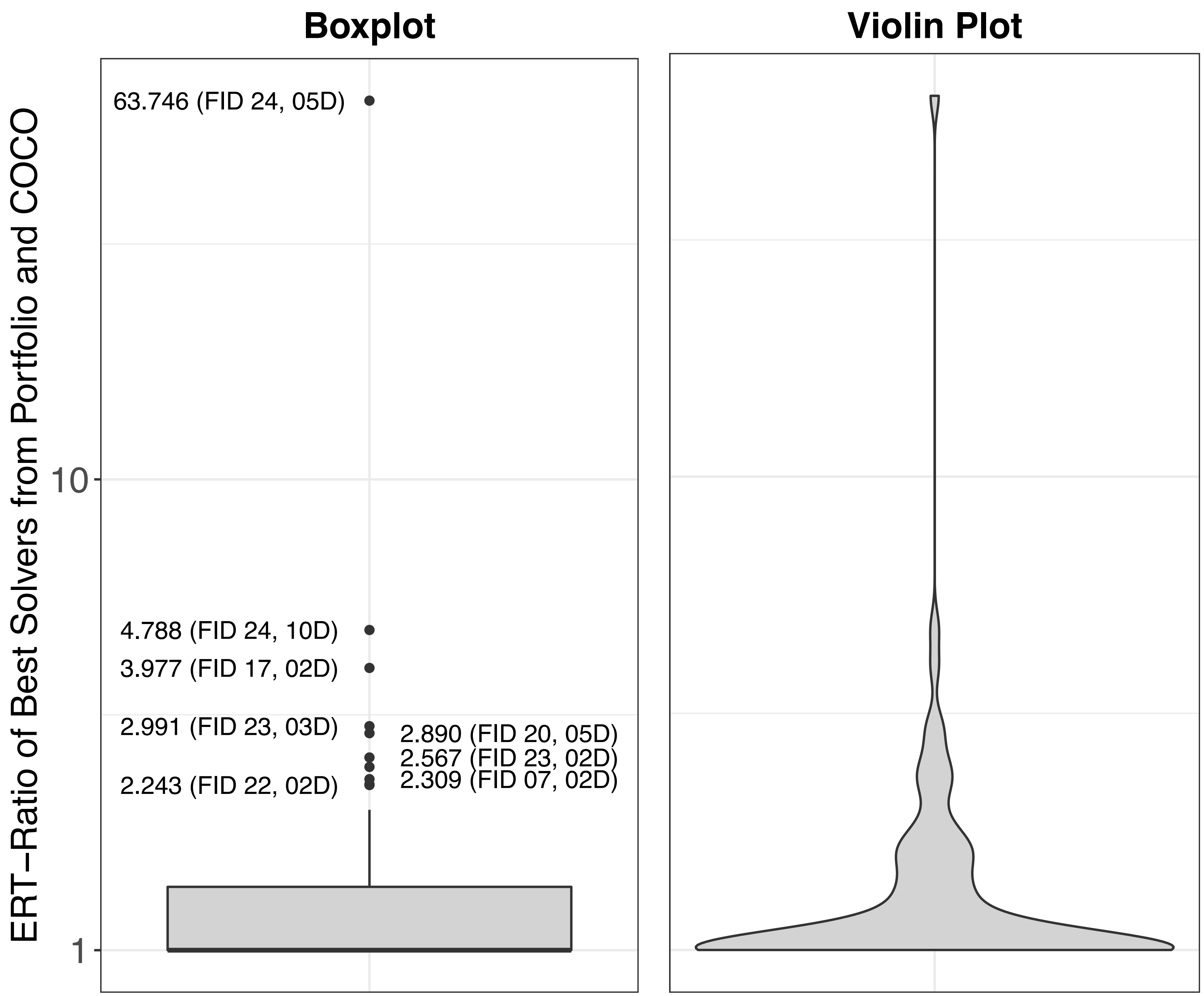
- algorithm portfolio:
  - ranked solvers per problem based on ERT
  - created 4 „sub-portfolios“ (1 per dimension) containing all solvers that ranked at least once in the Top 3
  - portfolio consists of all 12 solvers that were part of each „sub-portfolio“

Experimental Data

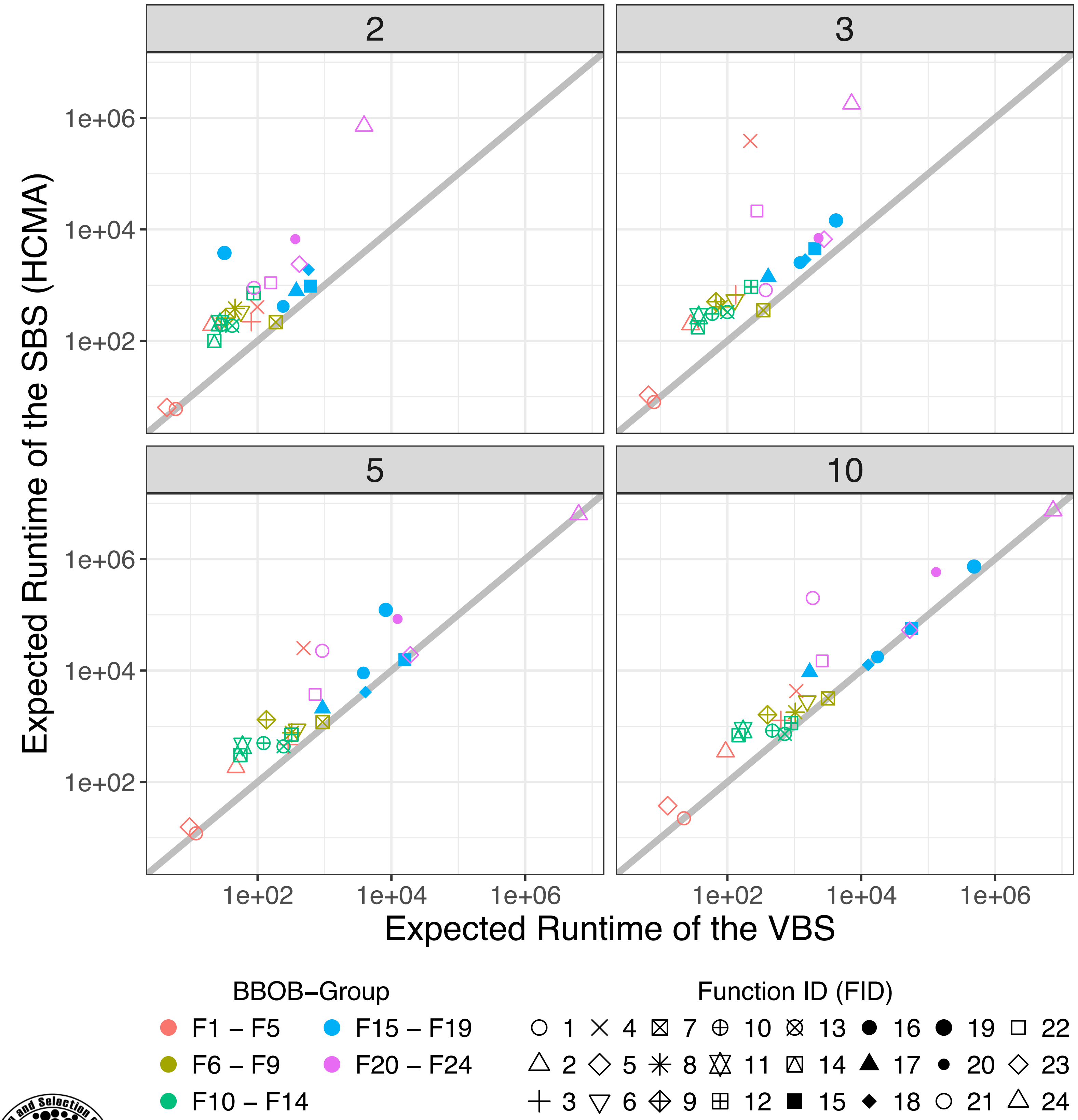
- **Deterministic Optimization Algorithms (2):**
  - BSrr and BSqi : variants of Brent-STEP
- **Multi-Level Approaches (5):**
  - MLSL: *Multi Level Single Linkage*
  - fmincon: variant of MLSL handling constrained nonlinear problems
  - fminunc: quasi-Newton variant of MLSL
  - HMLSL: hybrid version
  - MCS: *Multilevel Coordinate Search*
- **CMA-ES Variants (4):**
  - CMA-CSA: CMA-ES with cumulative step-size adaptation
  - IPOP400D: restart version with increasing population size
  - HCMA: hybrid CMA-ES, combining a BIPOP self-adaptive surrogate-assisted CMA-ES with STEP and NEWUOA

    Ilya Loshchilov, Marc Schoenauer, and Michèle Sebag. *Bi-Population CMA-ES Algorithms with Surrogate Models and Line Searches*. In: Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation (GECCO), pp. 1177 – 1184. ACM, July 2013.
  - SMAC-BBOB: sequential model-based algorithm configuration (SMAC) on BBOB
- **Others (1):**
  - OQNLP = OptQuest/NLP: a commercial, heuristic, multistart algorithm that was designed to find the global optima of smooth constrained nonlinear and mixed-integer programs

Exploratory Data Analysis



Exploratory Data Analysis



- HCMA was only algorithm to solve all 96 problems
- still often inferior to the remaining 11 solver
  - clearly potential for AS
- performance clearly depends on problem type

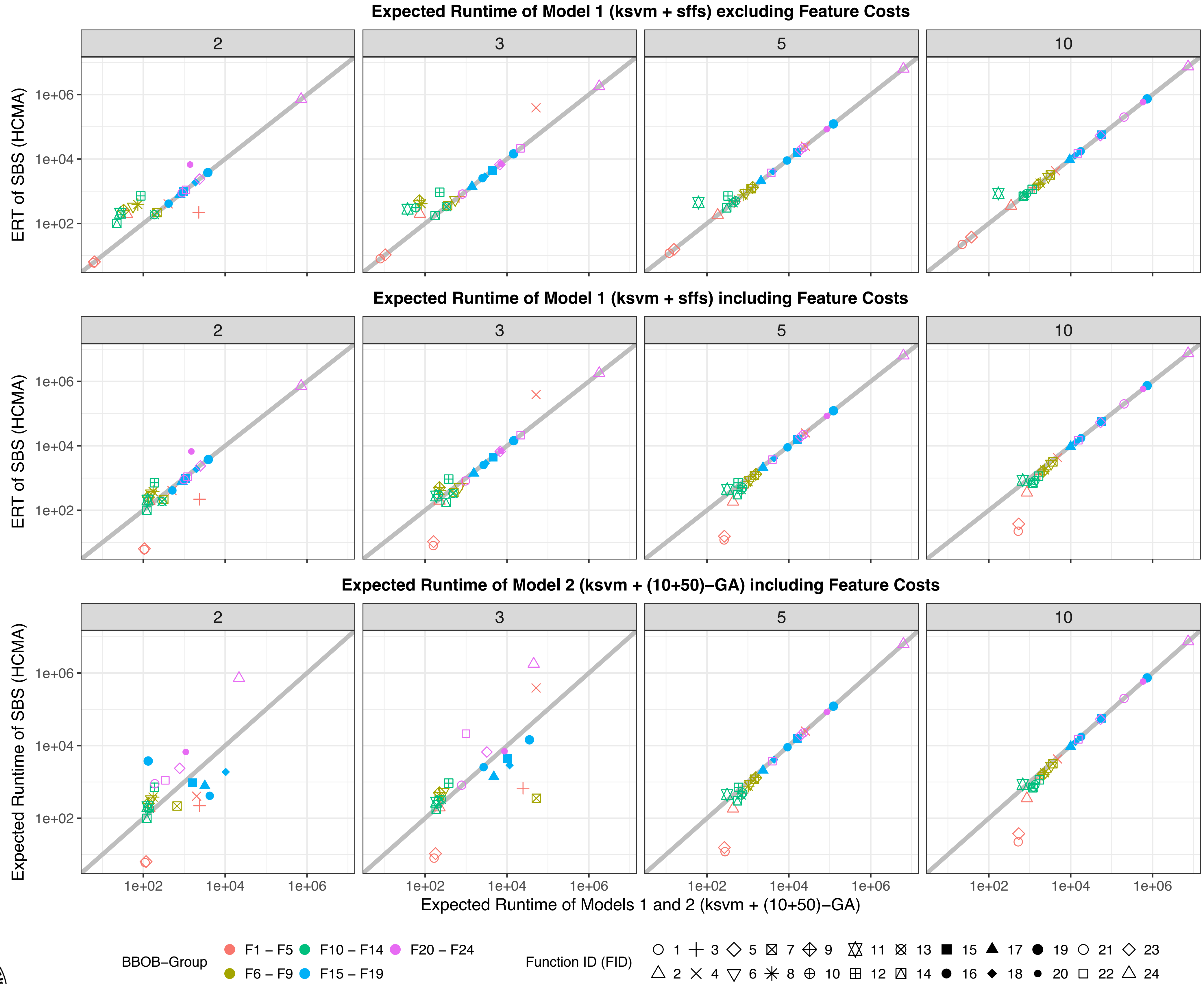
Algorithm Selection and Results

- **considered classification, regression and paired regression**
- **used random forests, trees, SVMs and gradient boosting**
  (mostly in their default configurations from R)
- **tried it without feature selection**
- **also ran experiments with four different feature selection strategies: sffs, sfbs, (10+5)-GA and (10+50)-GA**
- **trained algorithm selectors with the R-package mlR**
  Bernd Bischl, Michel Lang, Lars Kotthoff, Julia Schiffner, Jakob Richter, Erich Studerus, Giuseppe Casalicchio, and Zachary M. Jones. *mlr: Machine Learning in R*. In: Journal of Machine Learning Research (JMLR) , 17(170): pp. 1 - 5, 2016.
- **assessed performance of all models with leave-one-function-out crossvalidation**
- **best algorithm selectors were classification-based SVMs**
  (using different feature selection strategies)

Algorithm Selection and Results

Dim	BBOB-Group	Relative Expected Runtime of the 12 Solvers from the Considered Algorithm Portfolio and the 2 Presented Algorithm Selection-Models													
		BSqi	BSrr	CMA-CSA	fmincon	fminunc	HCMA	HMLSL	IPOP-400D	MCS	MLSL	OQNLP	SMAC-BBOB	AS-Model # 1	AS-Model # 2
2	F1 - F5	**1.2**	1.3	54.8	11.0	11.8	3.7	14.6	18.4	5.8	15.5	17.0	22 014.9	16.6	20.3
	F6 - F9	18 516.7	9 708.2	7.4	18.6	19.2	5.8	1.7	5.7	11.3	24.2	**1.5**	27 518.6	3.1	3.5
	F10 - F14	7 649.2	7 481.5	8.3	**1.0**	62.7	6.3	**1.0**	10.7	322.7	**1.0**	4.9	29 353.2	4.7	4.0
	F15 - F19	7 406.6	14 710.3	14.7	7 392.0	7 367.7	25.3	8.1	15.5	**7.7**	7 391.7	7 351.2	29 354.8	26.2	10.1
	F20 - F24	84.8	14 768.5	7 351.9	4.1	14.5	44.9	3.9	14 679.3	11.4	**2.1**	2.7	22 014.6	42.5	3.0
	all	6 240.7	9 318.4	1 549.1	1 546.5	1 556.7	17.7	**6.0**	3 068.4	74.3	1 547.9	1 536.9	25 990.1	19.3	8.4
3	F1 - F5	**1.3**	1.3	7 367.9	85.2	132.1	356.1	6.8	14 686.6	45.9	55.9	7 347.6	22 015.1	58.4	94.9
	F6 - F9	331.2	9 527.4	4.7	38.5	9 173.7	4.5	**1.9**	6.5	31.4	9 173.4	2.5	36 690.3	3.3	39.9
	F10 - F14	29 356.3	14 712.1	8.9	**1.0**	4.1	5.0	1.0	12.3	8 132.7	1.0	9.3	29 353.4	4.8	3.6
	F15 - F19	14 698.2	22 026.2	**1.6**	14 701.2	14 699.5	2.6	11.4	7 339.4	7 346.9	14 700.0	14 686.2	36 690.3	2.8	7.1
	F20 - F24	14 741.8	14 758.7	7 389.4	7 339.6	14 677.4	66.8	2.3	22 015.1	7 342.4	7 339.8	**1.9**	22 014.8	67.0	3.4
	all	12 304.7	12 316.7	3 077.4	4 616.2	7 677.5	90.4	**4.8**	9 178.9	4 769.4	6 132.4	4 593.1	29 047.1	28.3	29.4
5	F1 - F5	**1.4**	1.4	7 533.6	14 678.4	14 679.2	12.0	17.5	14 688.7	14 678.1	14 678.5	14 678.0	22 015.1	22.7	22.9
	F6 - F9	27 597.4	36 690.3	5.6	9 173.5	9 173.8	3.9	**2.4**	4.9	28.8	9 173.4	9 173.5	36 690.3	4.8	4.8
	F10 - F14	22 032.8	29 360.3	8.9	1.0	11.9	4.2	**1.0**	13.6	22 019.2	**1.0**	10.7	36 690.3	5.2	5.2
	F15 - F19	36 690.3	36 690.3	**3.1**	36 690.3	36 690.3	4.3	7 346.1	29 352.5	36 690.3	36 690.3	29 352.5	36 690.3	4.4	4.4
	F20 - F24	22 053.6	22 050.8	7 400.0	14 678.9	22 014.9	**7.7**	7 339.8	22 017.4	14 681.0	22 015.0	14 676.8	22 014.9	7.8	7.8
	all	21 428.3	24 469.8	3 114.6	15 289.0	16 819.9	**6.5**	3 063.8	13 765.8	18 352.4	16 817.4	13 761.8	30 575.6	9.1	9.2
10	F1 - F5	**1.6**	1.6	14 691.0	14 679.9	14 682.7	2.7	7 365.5	14 698.8	14 680.0	14 679.9	14 678.3	22 015.7	16.3	16.3
	F6 - F9	36 690.3	27 563.9	4.3	9 173.4	9 173.8	**2.2**	4.1	9 181.9	9 188.1	9 173.4	9 173.9	36 690.3	2.7	2.7
	F10 - F14	29 359.3	29 359.8	8.4	1.1	15.4	2.8	**1.1**	7 352.5	22 018.7	**1.1**	12.0	36 690.3	3.7	3.7
	F15 - F19	36 690.3	36 690.3	**1.7**	36 690.3	36 690.3	2.0	22 028.5	29 352.5	36 690.3	36 690.3	36 690.3	36 690.3	2.1	2.1
	F20 - F24	36 690.3	29 367.0	14 685.9	22 015.2	22 015.0	**23.6**	14 677.1	29 352.8	22 018.9	22 014.6	22 014.9	36 690.3	23.7	23.7
	all	27 519.5	24 472.9	6 123.0	16 817.8	16 821.3	**6.9**	9 182.4	18 354.6	21 408.0	16 817.6	16 819.7	33 633.1	10.0	10.0
all	F1 - F5	**1.4**	1.4	7 411.8	7 363.6	7 376.5	93.6	1 851.1	11 023.1	7 352.4	7 357.4	9 180.2	22 015.2	28.5	38.6
	F6 - F9	20 783.9	20 872.4	5.5	4 601.0	6 885.1	4.1	**2.5**	2 299.8	2 314.9	6 886.1	4 587.9	34 397.4	3.5	12.7
	F10 - F14	22 099.4	20 228.4	8.7	1.0	23.5	4.6	**1.0**	1 847.3	13 123.3	**1.0**	9.3	33 021.8	4.6	4.1
	F15 - F19	23 871.3	27 529.3	**5.2**	23 868.5	23 861.9	8.6	7 348.5	16 515.0	20 183.8	23 868.1	22 020.0	34 856.4	8.9	5.9
	F20 - F24	18 392.6	20 236.3	9 206.8	11 009.4	14 680.5	**35.8**	5 505.8	22 016.2	11 013.4	12 842.9	9 174.1	25 683.7	35.3	***9.5***
	all	16 873.3	17 644.5	3 466.0	9 567.4	10 718.9	**30.4**	3 064.3	11 091.9	11 151.0	10 328.8	9 177.9	29 811.5	16.7	***14.2***

Algorithm Selection and Results



Algorithm Selection and Results

only used by AS #1

- **2x y-distribution**
  (kurtosis, # peaks)
- **2x information content**
  (max. information content, settling sensitivity)
- **1x basic**
  (best objective value within the sample)

used by both models

- **1x levelset**
  (MMCE-ratio of LDA & MDA)
- **1x cell mapping**
  (STD of distances between center and worst observation)
- **1x y-distribution**
  (skewness)

only used by AS #2

- **2x meta model**
  (smallest abs., non-intercept coefficient of linear model, adj. R^2 of quadratic model)
- **4x nearest better clustering**
  (ratios of STDs, ratios of arithmetic means, correlation, indegree)

Takeaway

- **first extensive AS study in the field of *continuous* optimization**
  (thanks to previous works, e.g., BBOB & COCO, low-budget ELA, flacco and mlr)
- **showed that the combination of ELA and ML can be very powerful**
  (we reduced the mean relative ERT of the SBS by half using a small set of features)
- **this work has recently been submitted to TEVC:**
  „Pascal Kerschke and Heike Trautmann. *Automated Algorithm Selection on Continuous Black-Box Problems By Combining Exploratory Landscape Analysis and Machine Learning.* In: IEEE Transactions on Evolutionary Computation. IEEE (currently under review).“

Special Issue on

„Algorithm Selection and Configuration in Evolutionary Computation"

— Submission Deadline: November 30, 2017 —

JOURNAL:

- Evolutionary Computation Journal
- MIT Press (http://ecj.napier.ac.uk)

POSSIBLE TOPICS (not limited to those):

- automated algorithm selection
- specific machine learning concepts
- configuration methods
- performance analysis
- features and diversity problem instances
- benchmarking concepts
- exploratory landscape analysis

Frank Neumann
University of Adelaide
(Associate Editor)

Holger H. Hoos
Universiteit Leiden
(Guest Editor)

Heike Trautmann
University of Münster
(Guest Editor)

Authors should submit their manuscripts to the Evolutionary Computation Editorial Manager at http://ecj.napier.ac.uk. When submitting a paper, please send at the same time an email to Frank Neumann (frank.neumann@adelaide.edu.au) and a copy to ecj@napier.ac.uk mentioning the special issue, the paper id, title, and author list to inform us about the submission.