

# R Notebook

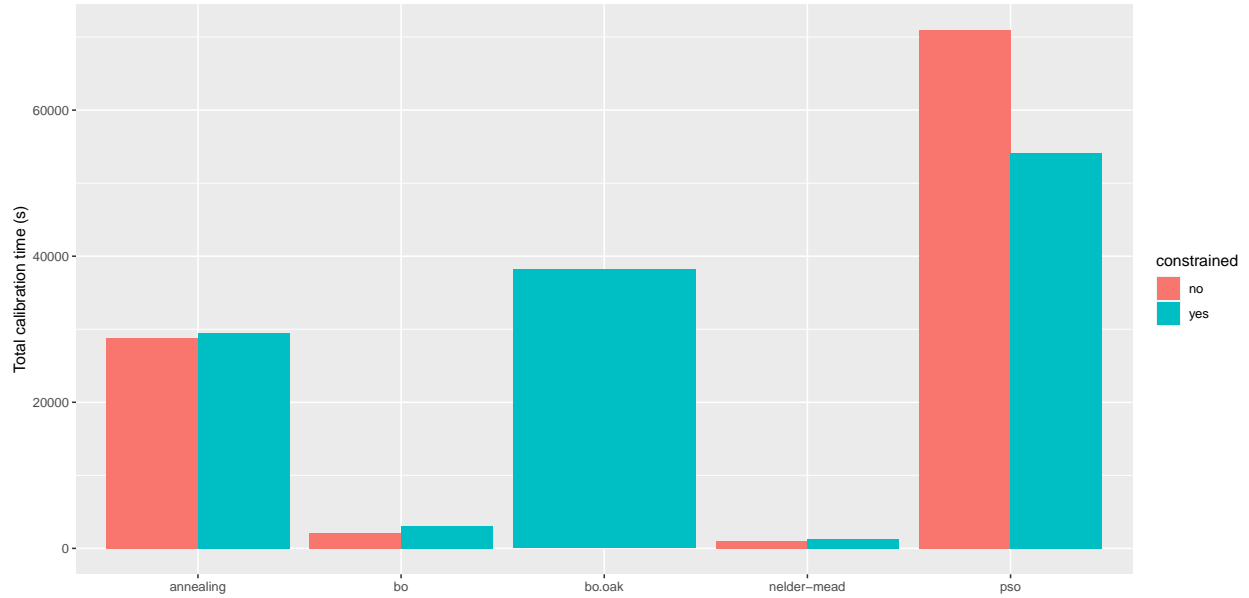
## Stepwise calibration of lung cancer model with constraints

- 9 age groups, 11 parameters per age group: total of 99 parameters to be calibrated.
- 9 of those parameters (one per age group) are probabilities of developing cancer, with the constraint of these probabilities increasing with age.

method	group	time	iterations	error	delay	constrained
nelder-mead	1	86.21	839	0.1685787	0.100	yes
nelder-mead	2	79.72	775	0.1685787	0.100	yes
nelder-mead	3	81.66	793	0.1685787	0.100	yes
nelder-mead	4	101.61	985	0.2112280	0.100	yes
nelder-mead	5	128.52	1236	0.6645467	0.100	yes
nelder-mead	6	127.03	1233	1.9795078	0.100	yes
nelder-mead	7	231.11	2242	77.2779470	0.100	yes
nelder-mead	8	201.02	1948	81.3999553	0.100	yes
nelder-mead	9	288.74	2796	83.0570106	0.100	yes
bo	1	119.78	40	0.1237346	0.100	yes
bo	2	123.29	40	0.1284176	0.100	yes
bo	3	176.05	40	0.1537532	0.100	yes
bo	4	355.28	40	0.2191810	0.100	yes
bo	5	591.22	40	0.7689339	0.100	yes
bo	6	485.97	40	2.2675644	0.100	yes
bo	7	373.80	40	76.7939092	0.100	yes
bo	8	507.19	40	80.8674985	0.100	yes
bo	9	357.18	40	83.0535954	0.100	yes
annealing	1	2343.80	22637	0.1137244	0.100	yes
annealing	2	2376.90	22925	0.1137280	0.100	yes
annealing	3	2323.23	22361	0.1137493	0.100	yes
annealing	4	2586.64	24893	0.1562925	0.100	yes
annealing	5	3512.64	33701	0.6084468	0.100	yes
annealing	6	3763.44	36113	1.9237158	0.100	yes
annealing	7	3085.71	29561	77.1249233	0.100	yes
annealing	8	5591.71	53501	81.2325550	0.100	yes
annealing	9	3917.08	37421	82.8610077	0.100	yes
pso	1	1757.15	16980	0.1126824	0.100	yes
pso	2	964.73	9298	0.1126824	0.100	yes
pso	3	1039.33	10016	0.1126824	0.100	yes
pso	4	1121.64	10784	0.1552969	0.100	yes
pso	5	1642.14	15786	0.6158138	0.100	yes
pso	6	1735.28	16691	1.9305617	0.100	yes
pso	7	21155.12	202950	77.1683602	0.100	yes
pso	8	21197.11	202950	81.3903838	0.100	yes
pso	9	3465.25	33201	83.0449091	0.100	yes
nelder-mead	1	86.94	839	0.1685787	0.100	no
nelder-mead	2	80.58	775	0.1685787	0.100	no

method	group	time	iterations	error	delay	constrained
nelder-mead	3	73.84	708	0.1991849	0.100	no
nelder-mead	4	89.06	853	0.2418408	0.100	no
nelder-mead	5	163.21	1566	0.6940669	0.100	no
nelder-mead	6	135.69	1298	2.0092999	0.100	no
nelder-mead	7	74.80	716	76.1249879	0.100	no
nelder-mead	8	141.97	1356	80.0057314	0.100	no
nelder-mead	9	139.86	1336	81.2596918	0.100	no
annealing	1	2349.66	22661	0.1128548	0.100	no
annealing	2	2329.76	22433	0.1128691	0.100	no
annealing	3	2341.87	22541	0.1128693	0.100	no
annealing	4	2739.82	26357	0.1554071	0.100	no
annealing	5	3572.21	34373	0.6075151	0.100	no
annealing	6	3849.34	36965	1.9233890	0.100	no
annealing	7	2858.78	27329	76.0270601	0.100	no
annealing	8	4479.59	42893	79.7575312	0.100	no
annealing	9	4288.06	41141	81.0635058	0.100	no
bo	1	71.93	40	0.1237346	0.100	no
bo	2	69.49	40	0.1301805	0.100	no
bo	3	92.83	40	0.1840909	0.100	no
bo	4	190.90	40	0.2901077	0.100	no
bo	5	469.89	40	0.9046652	0.100	no
bo	6	388.85	40	2.4350920	0.100	no
bo	7	260.54	40	74.8822406	0.100	no
bo	8	320.04	40	80.7263702	0.100	no
bo	9	218.78	40	83.2432799	0.100	no
pso	1	1701.46	16418	0.1126843	0.100	no
pso	2	1096.36	10571	0.1126843	0.100	no
pso	3	1092.46	10499	0.1126843	0.100	no
pso	4	1040.99	10020	0.1552424	0.100	no
pso	5	1409.97	13557	0.6086457	0.100	no
pso	6	1111.77	10680	1.9245714	0.100	no
pso	7	21146.37	202950	76.0244206	0.100	no
pso	8	21096.04	202950	79.7500776	0.100	no
pso	9	21244.08	202950	81.0407388	0.100	no
bo.oak	1	5609.74	25	0.0011802	0.001	yes
bo.oak	2	2928.64	25	0.0363753	0.001	yes
bo.oak	3	3407.21	25	0.1808856	0.001	yes
bo.oak	4	5975.08	25	0.1983608	0.001	yes
bo.oak	5	4338.44	25	0.6312610	0.001	yes
bo.oak	6	3127.86	25	2.5961679	0.001	yes
bo.oak	7	4822.83	25	63.4569127	0.001	yes
bo.oak	8	4260.71	25	63.4577852	0.001	yes
bo.oak	9	3703.87	25	65.1182209	0.001	yes

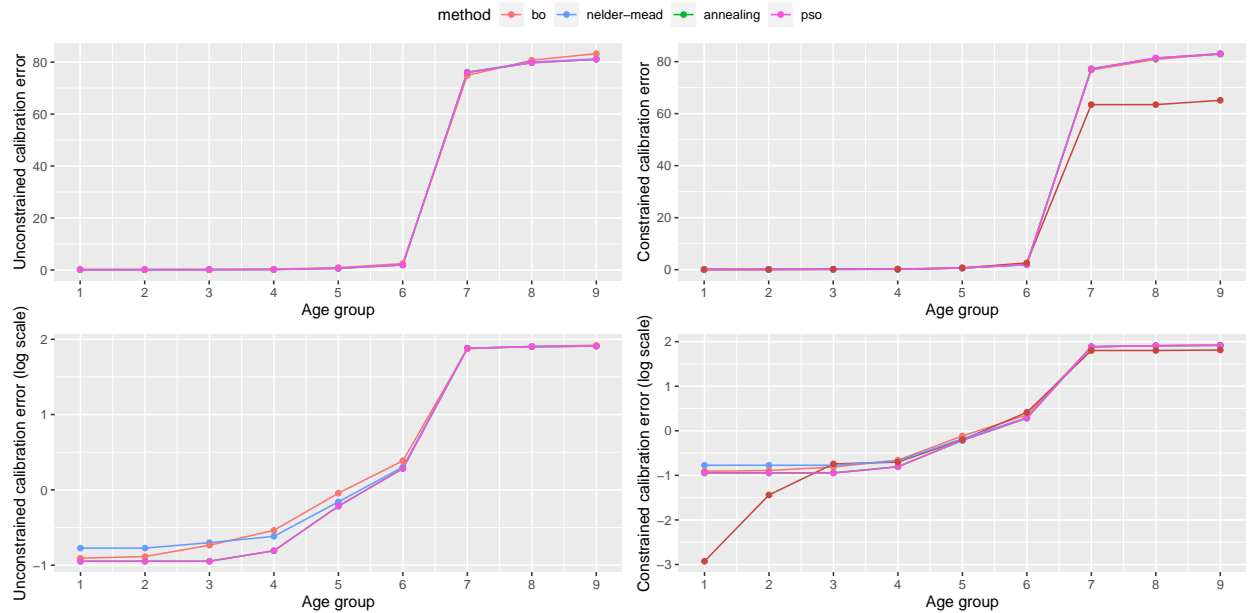
## Overall results



With a delay of 100 ms (therefore  $t_{sim} \approx 100\text{ms}$ ), both PSO and SA need a lot of time due to its large number of iterations. NM and BO use a comparable amount of time for similar results, even if for this simulation time NM is faster. Constrained calibration takes more time for all methods except PSO, but seems to be reasonable for these dimensionalities.

In absolute terms, calibration with BO using the stepwise method takes around **50 minutes**. Even the unconstrained version of the regular calibration proved to be unfeasible, requiring many days of computation while finding low-quality solutions. Adding constraints to the regular calibration would increase even further the computational cost of an already unfeasible problem.

## Calibration errors by age group



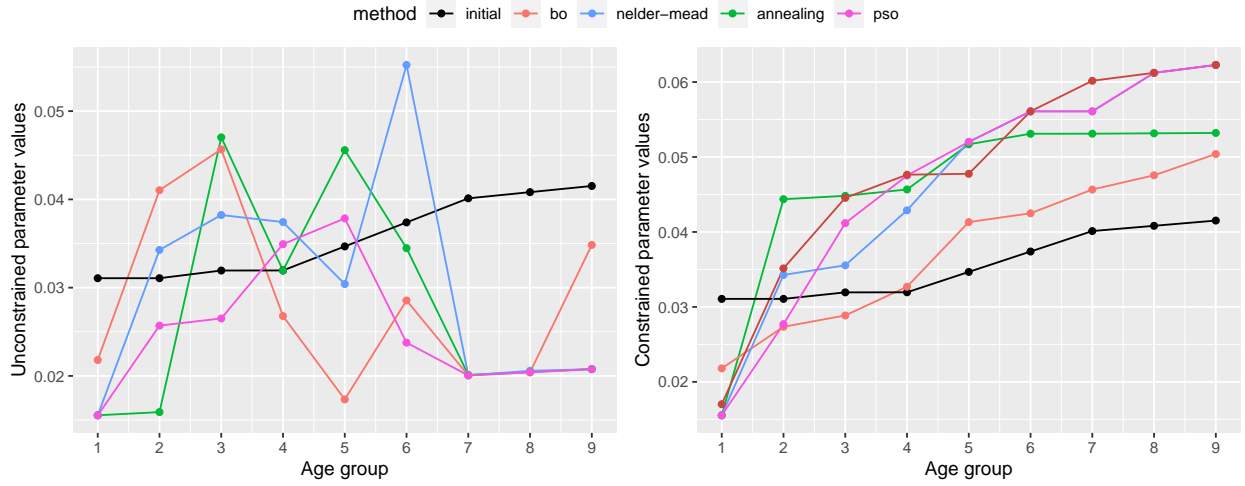
Solutions reached by the different methods have very similar errors (unconstrained calibration on the left,

constrained calibration on the right), so the final results are comparable among methods. The slight increase in BO error at the final age group might not be too important due to how these models work (it is very small and the outputs at the end of the simulation have less impact).

### Calibrated parameters by age group

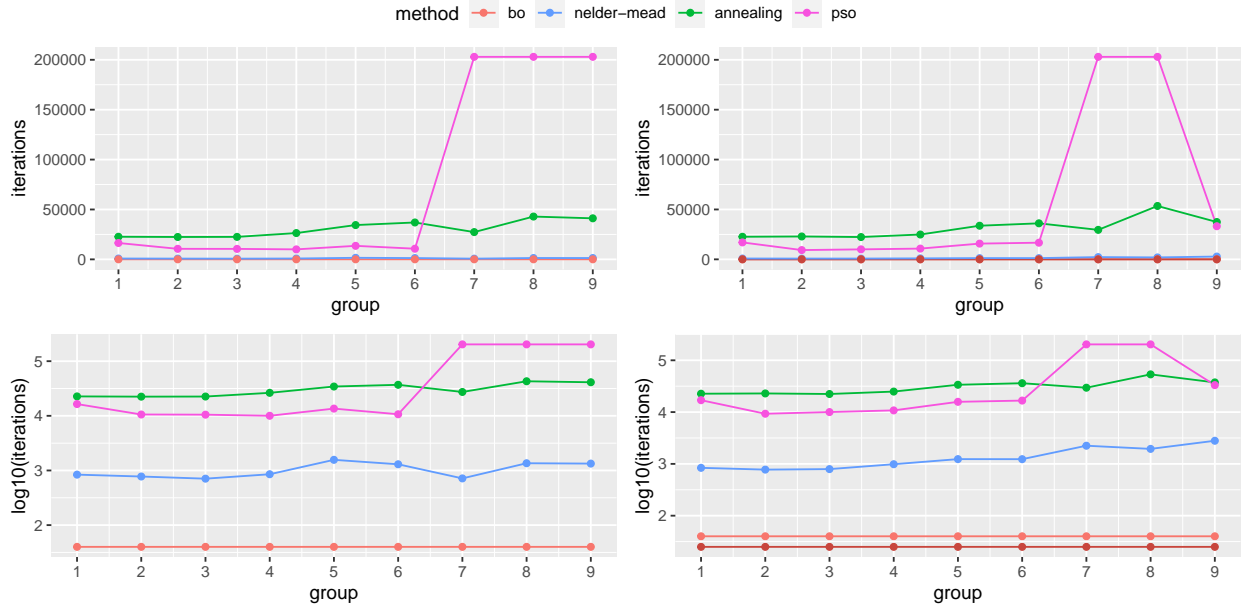
	method	param	value	constrained	error	distance
1	initial	1	0.03107	yes	0.00000	0.00000
2	initial	2	0.03107	yes	0.00000	0.00000
3	initial	3	0.03194	yes	0.00000	0.00000
4	initial	4	0.03195	yes	0.00000	0.00000
5	initial	5	0.03467	yes	0.00000	0.00000
6	initial	6	0.03739	yes	0.00000	0.00000
7	initial	7	0.04012	yes	0.00000	0.00000
8	initial	8	0.04082	yes	0.00000	0.00000
9	initial	9	0.04152	yes	0.00000	0.00000
10	nelder-mead	1	0.01554	yes	-0.01553	0.01553
11	nelder-mead	2	0.03427	yes	0.00320	0.00320
12	nelder-mead	3	0.03555	yes	0.00361	0.00361
13	nelder-mead	4	0.04289	yes	0.01094	0.01094
14	nelder-mead	5	0.05201	yes	0.01734	0.01734
15	nelder-mead	6	0.05609	yes	0.01870	0.01870
16	nelder-mead	7	0.05609	yes	0.01597	0.01597
17	nelder-mead	8	0.06123	yes	0.02041	0.02041
18	nelder-mead	9	0.06228	yes	0.02076	0.02076
19	bo	1	0.02180	yes	-0.00927	0.00927
20	bo	2	0.02735	yes	-0.00372	0.00372
21	bo	3	0.02886	yes	-0.00308	0.00308
22	bo	4	0.03269	yes	0.00074	0.00074
23	bo	5	0.04132	yes	0.00665	0.00665
24	bo	6	0.04249	yes	0.00510	0.00510
25	bo	7	0.04566	yes	0.00554	0.00554
26	bo	8	0.04756	yes	0.00674	0.00674
27	bo	9	0.05040	yes	0.00888	0.00888
28	annealing	1	0.01554	yes	-0.01553	0.01553
29	annealing	2	0.04437	yes	0.01330	0.01330
30	annealing	3	0.04482	yes	0.01288	0.01288
31	annealing	4	0.04567	yes	0.01372	0.01372
32	annealing	5	0.05170	yes	0.01703	0.01703
33	annealing	6	0.05310	yes	0.01571	0.01571
34	annealing	7	0.05311	yes	0.01299	0.01299
35	annealing	8	0.05316	yes	0.01234	0.01234
36	annealing	9	0.05321	yes	0.01169	0.01169
37	pso	1	0.01553	yes	-0.01554	0.01554
38	pso	2	0.02770	yes	-0.00337	0.00337
39	pso	3	0.04118	yes	0.00924	0.00924
40	pso	4	0.04752	yes	0.01557	0.01557
41	pso	5	0.05201	yes	0.01734	0.01734
42	pso	6	0.05609	yes	0.01870	0.01870
43	pso	7	0.05609	yes	0.01597	0.01597
44	pso	8	0.06122	yes	0.02040	0.02040
45	pso	9	0.06228	yes	0.02076	0.02076
91	bo.oak	1	0.01702	yes	-0.01405	0.01405

	method	param	value	constrained	error	distance
92	bo.oak	2	0.03514	yes	0.00407	0.00407
93	bo.oak	3	0.04455	yes	0.01261	0.01261
94	bo.oak	4	0.04764	yes	0.01569	0.01569
95	bo.oak	5	0.04776	yes	0.01309	0.01309
96	bo.oak	6	0.05609	yes	0.01870	0.01870
97	bo.oak	7	0.06017	yes	0.02005	0.02005
98	bo.oak	8	0.06123	yes	0.02041	0.02041
99	bo.oak	9	0.06228	yes	0.02076	0.02076



BO finds the solution that resembles the initial parameters the most. For some reason, Nelder-Mead and PSO find very similar constrained solutions for group ages 5+ and they can't be distinguished in the right figure.

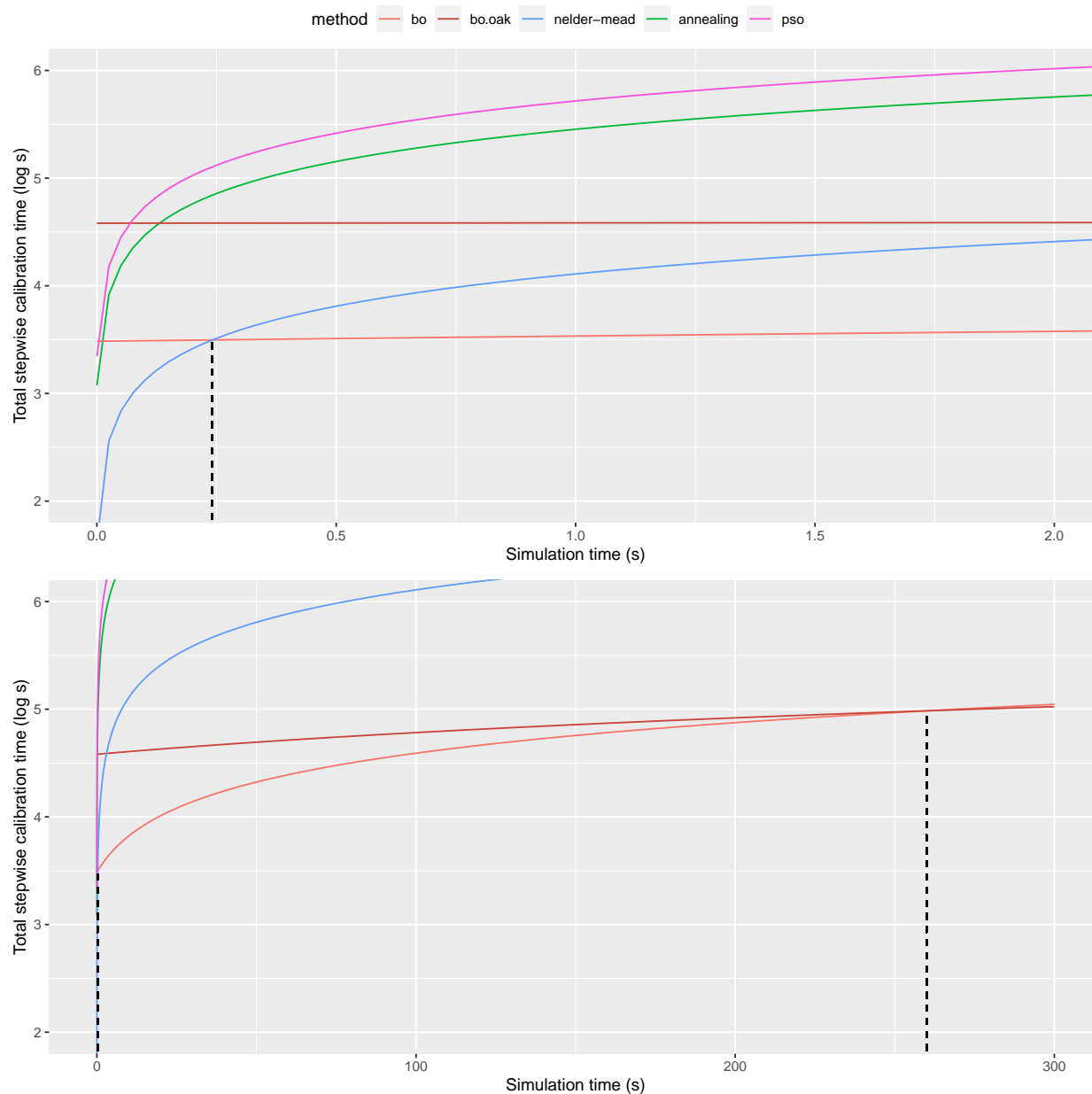
## Number of iterations per age group



BO was run with a fixed budget of 40 iterations, including 10 random initial observations. Even though SA

has a lower error we can see that it requires a very large number of iterations for a very modest improvement on the solution. PSO gets stuck in the last age groups and stops at the established maximum iterations.

## Critical simulation time for constrained stepwise calibration



The critical simulation time for the whole model with all 9 age groups (that **we projected at 300 seconds using conventional calibration**) is reduced to **~0.24 seconds using constrained stepwise calibration**. This proves that this method dramatically improves the efficiency of Bayesian Optimization by exploiting a major flaw of BO (high dimensionality), reducing the effective dimension of the problem from one 99-parameter problem to nine 11-parameter problems.

The critical simulation time doesn't change significantly when considering an unconstrained calibration.

~260s