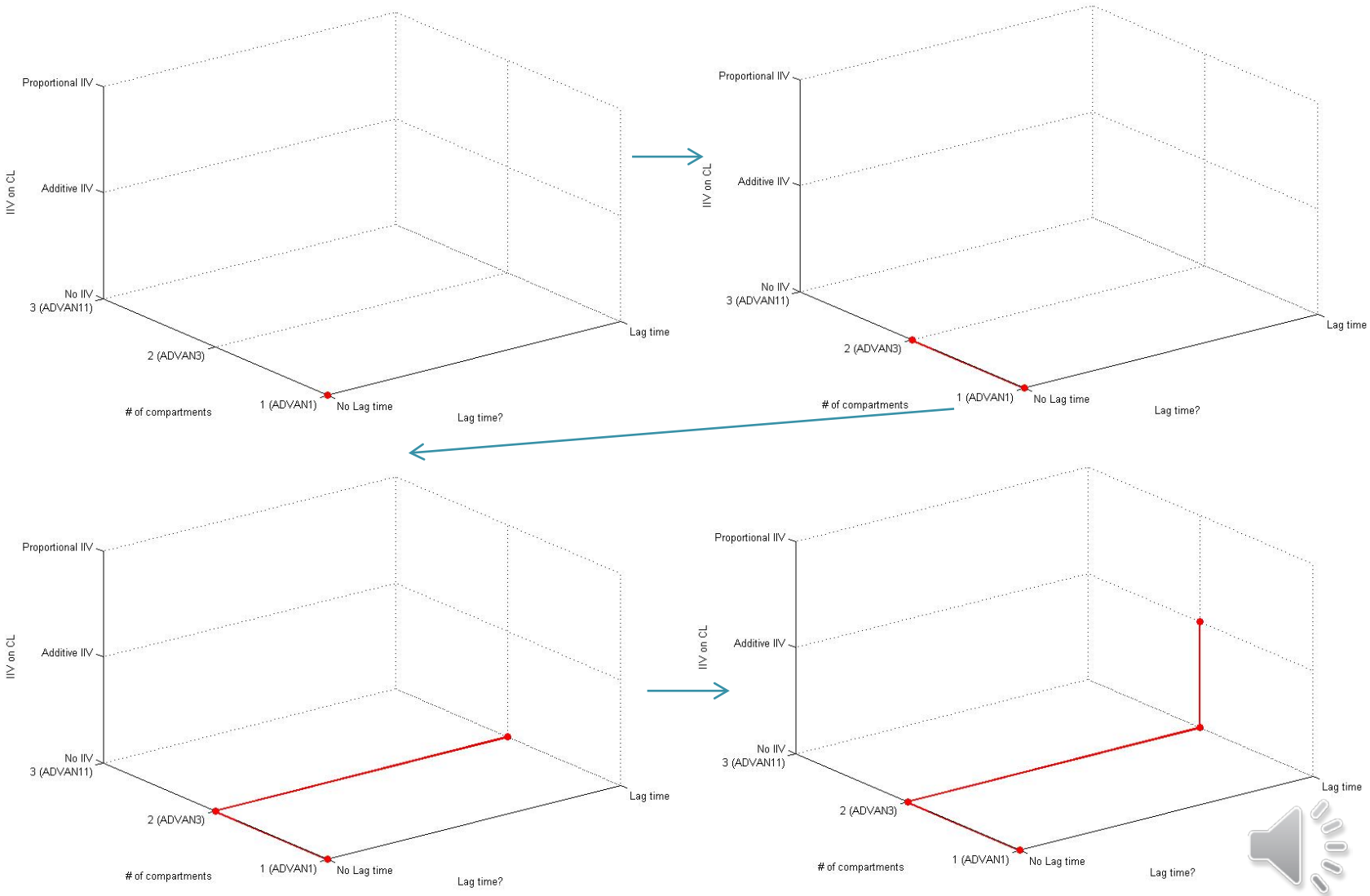
Machine search vs. humans in modeling
or
"I'm sorry Dave, I'm afraid I can't do that"

Mark Sale M.D.
Next Level Solutions

Nov 2, 2011

Current (local) search algorithm – assumption is path/starting point independence



Prior knowledge (Bias? Prejudice?) in model selection

- Certainly we know that weight is a predictor of volume, and the astrological sign could never predict anything
- So, never test V = F(wt), just put it in.
- And, never test (V = f(astrological sign), even if it shows up, you wouldn't believe it.

Right??

Astrological sign is a useful predictor of hockey success in Canada (Capricorn and Aries are the most likely to succeed) and football success in UK, and success in a number of sports.

Musch J, Grondin S. "Unequal competition as an Impediment to personal development: A review of relative age effect in sport", Developmental Review 21(2) 2001 147-167

The case for a global search (as opposed to local search)

- Unexpected things happen
  - Things we're sure about turn out not to be the case
  - Things we're sure cannot be turn out to be the case
  - You won't find either unless you look.
- Local minima problem

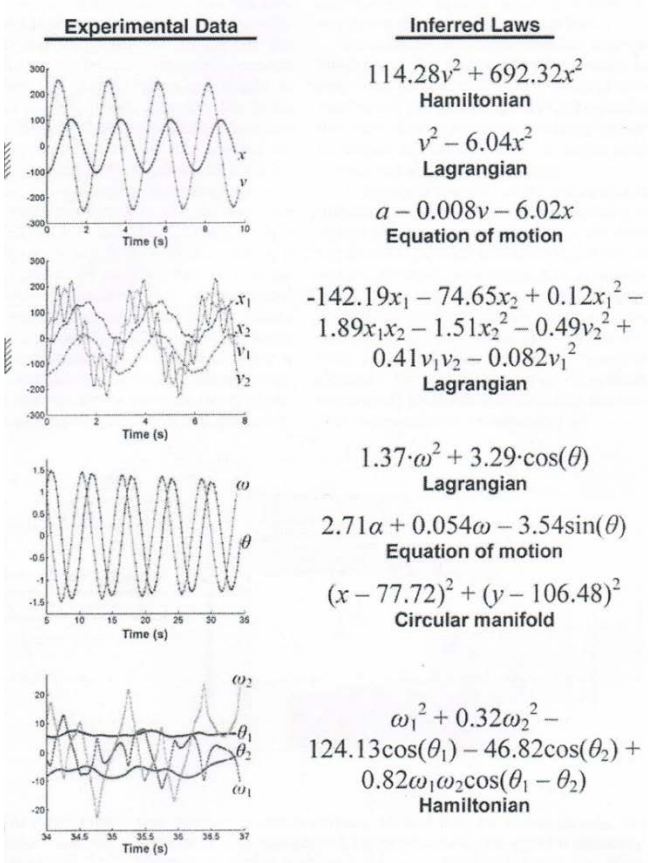
Another search algorithm: Genetic algorithm

- Reproduction of evolution/mutation/cross over/survival of the fittest.
- Widely used to optimize engineering systems.

But, creating "learning/understanding/insight" is different from "optimizing"- isn't it?

- "Distilling free-form natural laws from experimental data". Science 324: 81-85

- Used GA to find combination of elementary math function (+,-,*,sin, tan, ln) and data to derive equation for motion of double pendulum.



- Is learning/generating understanding/insight frequently just assembling existing pieces in new, useful (perhaps insightful) ways?

All science is either physics or stamp collecting* (Ernest Rutherford)

*And it appears that perhaps at least some physics is stamp collecting as well

I can't claim to have ever come up with anything genuinely novel. I'm a stamp collector

Similarly:

- **Automated refinement and inference of analytical models for metabolic networks**. Michael D Schmidt, Ravishankar R Vallabhajosyula, Jerry W Jenkins, Jonathan E Hood, Abhishek S Soni, John P Wikswo, Hod Lipson. *Physical Biology*, 2011; 8 (5): 055011 DOI:

Proposal:

- Hybrid modeling selection algorithm
- Combine robustness and efficiency of global search with biological understanding/experience, diagnostic plot evaluation and consideration of plausibility
- Avoid local minima in the search space by having a better starting point.
  - Start with global model search (which may require some iteration, using biological understanding etc.), to get a better starting point, more likely to be monotonically down.
  - The "search space" is user defined
  - The search criteria are user defined
  - Then forward addition/backward elimination for plausibility.

So, can we get in the ball park: Results of cross over trial of traditional vs. GA for 7 analyses from Sherer et. al. (test of just the search algorithm only)

Compound	**Final stepwise model**	**Final SOHGA model**	**$AIC_{SOHGA} - AIC_{stepwise}$ (% change)**
Citalopram, IV	BIC = 5760.2 AIC = 5,713.5 -2LL = 5,695.5	BIC = 5,436.2 AIC = 5,363.6 -2LL = 5,335.6	-357.9
DMAG, IV	BIC = 9,938.2 AIC = 9,862.5 -2LL = 9,832.5	BIC = 9,913.0 AIC = 9,847.4 -2LL = 9,821.4	-15.1
Escitalopram	BIC = 2,774.9 AIC = 2,729.1 -2LL = 2,707.1	BIC = 2,777.2 AIC = 2,735.6 -2LL = 2,715.6	6.5
Olanzapine, oral	BIC = 10,413.8 AIC = 10,365.8 -2LL = 10,347.8	BIC = 9,937.9 AIC = 9,895.3 -2LL = 9,879.3	-470.5
Perphenazine, oral	BIC = 601.1 AIC = 560.7 -2LL = 540.7	BIC = 604.4 AIC = 555.9 -2LL = 531.9	-4.8
Risperidone, oral	BIC = 5,188.5 AIC = 5,127.1 -2LL = 5,103.1	BIC = 4,824.7 AIC = 4,762.7 -2LL = 4,738.7	-364.4
Ziprasidone, oral	BIC = 4,880.8 AIC = 4,850.4 -2LL = 4,836.4	BIC = 4,759.4 AIC = 4,758.7 -2LL=4,746.7	-91.7

Comparison of traditional and GA final models final status (just GA, no final FA/BE step)

	Convergence			Covariance step	
	Final stepwise model	**Best SOHGA candidate**		**Final stepwise model**	**Best SOHGA candidate**
Citalopram, IV	Successful	Successful		Unsuccessful	Successful
DMAG, IV	Successful	Successful		Unsuccessful	Successful
Escitalopram	Successful	Successful		Successful	Successful
Olanzapine, oral	Successful after fixing K_a	Successful		Successful	Successful
Perphenazine, oral	Successful after fixing K_a	Successful		Unsuccessful	Successful
Risperidone, oral	Successful after fixing K_a	Successful		Successful	Successful
Ziprasidone, oral	Successful after fixing K_a	Successful		Successful	Successful

Comparison of traditional vs. GA final models:

- Identical structural models.
- The hybrid GA models included 50% (7 of 14) of significant covariates in the stepwise models and the stepwise model included 30% (7 of 23) of significant covariates in the final SOHGA models.
- SOHGA included fewer IIV terms
- So, more covariates, fewer IIV terms

Example:

- Motivated by a sponsors desire to have all decision point models with successful covariance step
- Could not find *any* models with successful covariance (1|2 compartment, ETA on CL, V, concomitant med on CL and F, lag time)
- So, we proposed that we start with global search algorithm to find a model with successful covariance step.

Outcome of real world example of hybrid GA/FABE modeling

- Start by generating the hypotheses (this part doesn't change)
- Found a starting point with successful covariance step (key was inter occasion variability in CL and initial estimates, along with con-med on CL)
- Final Model from SOHGA was a local minimum. Fixed by removing one OMEGA term.
- Final model (after FA/BE) was similar to GA model, mostly rearranging things.
- Also permitted searching on initial estimates, CTYPE, NUMPOINTS, ADVAN6|8|9|13 etc.

Advantages of hybrid GA/FABE

- Robustness and efficiency of global search
- Biological insight/evaluation of diagnostics and plausibility of FA/BE
- Faster (1000's of models by GA followed by dozens of model by hand, rather than 100's of models by hand).
- More objective, more thorough.

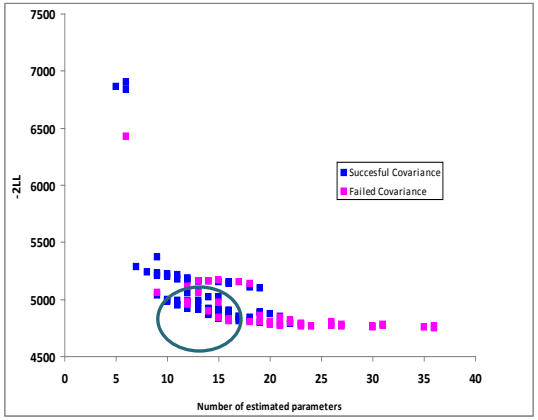
Multi objective optimization

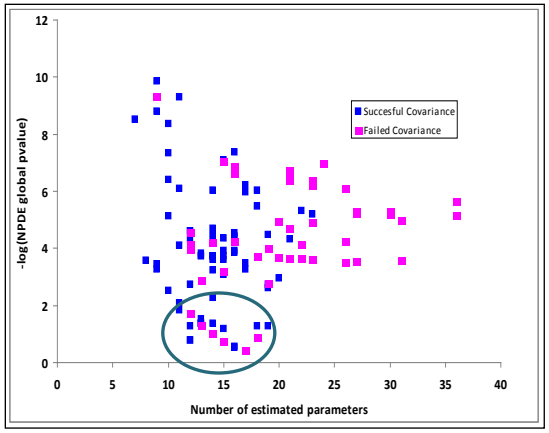
- Single objective GA uses a composite "fitness" function – combination of -2LL and other things (user defined penalties for parameters, convergence, covariance etc)
- This is pretty rigid and arbitrary – who's to say that an additional THETA is "worth" so many points (besides Akaike)
- People doing GA found that the decision maker didn't want to be told – here is the "best" option. The decision makers felt that certain "subjective, experience based" criteria couldn't be captured.

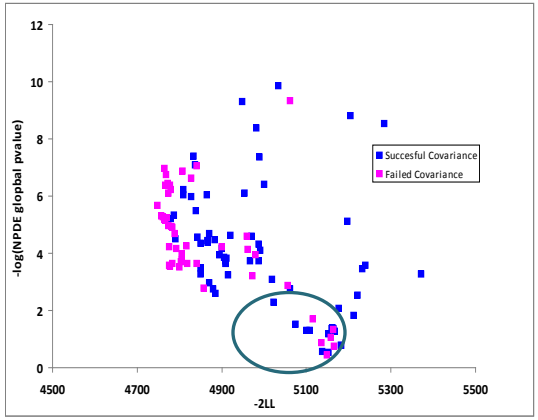
So, optimize over many criteria, present user with a variety of options

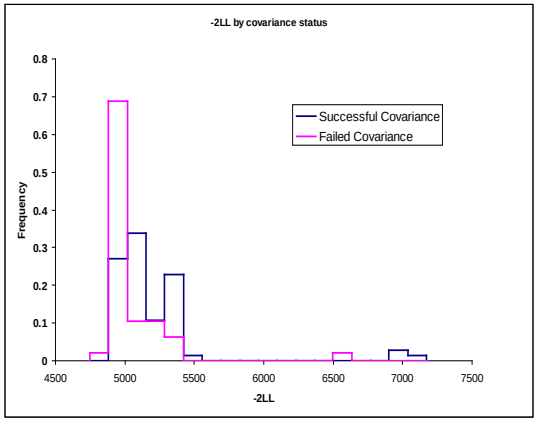
- So, you want a bridge built, or a model selected. There are trade offs that may be difficult to quantify
  - Bridge - More expensive – lasts longer
  - Model – Better -2LL/VPC, more parameters
- Present a variety of bridges/models, Some with more parameters, some with fewer, some with successful covariance etc.

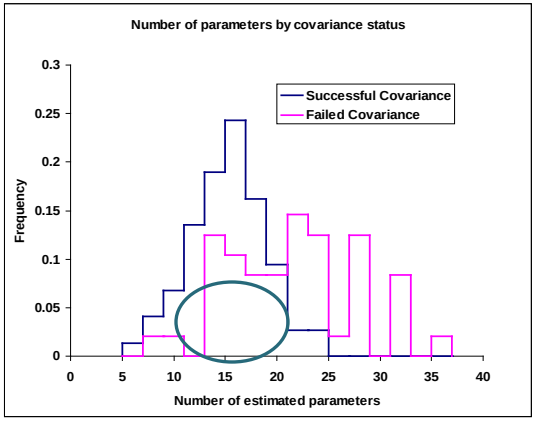
Results: Non dominated solutions

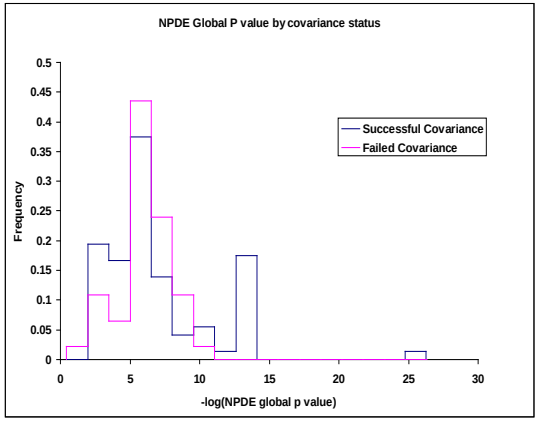




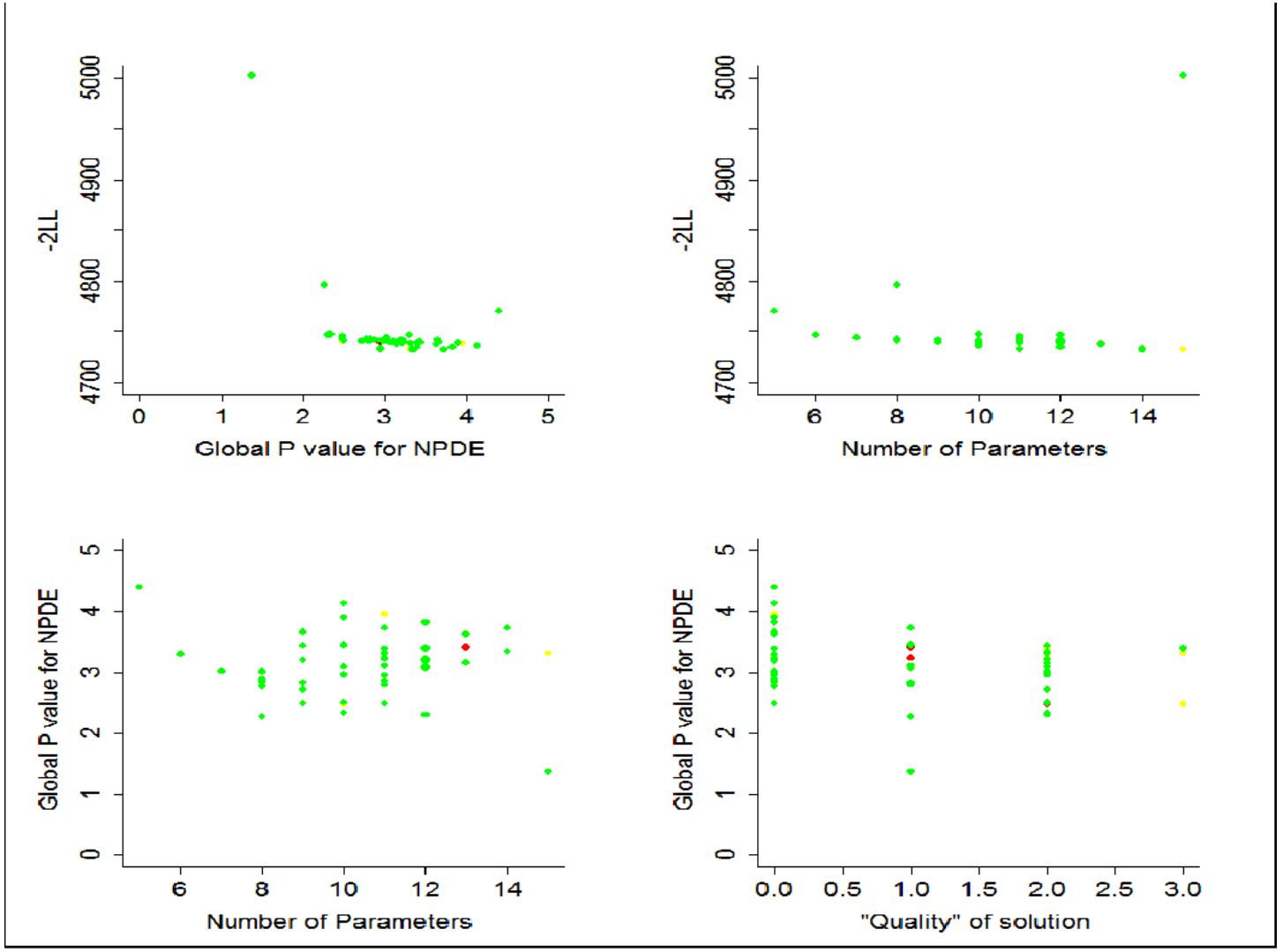




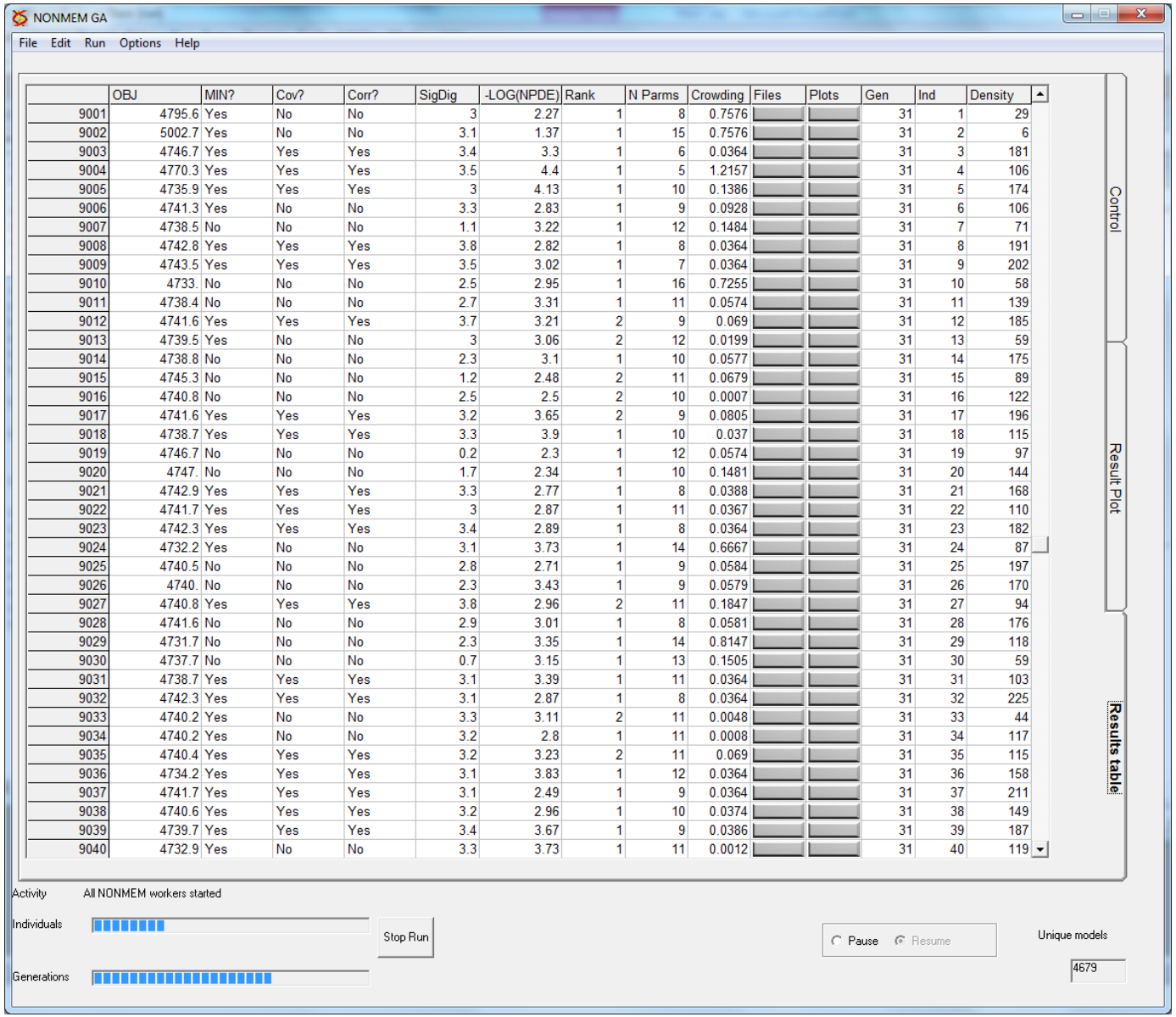




Another (simpler) example: Only covariates/OMEGA/SIGMA 31 generations

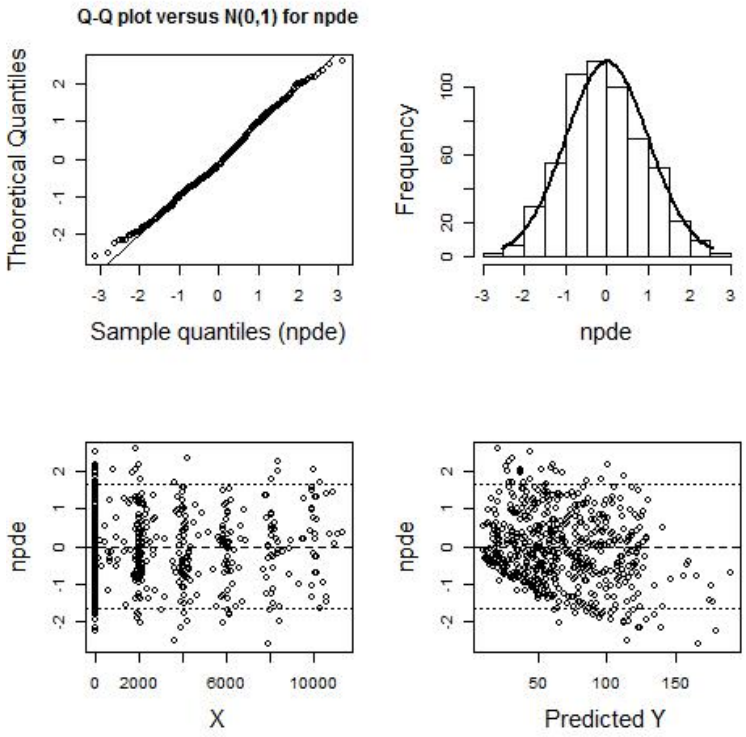


Trade offs:

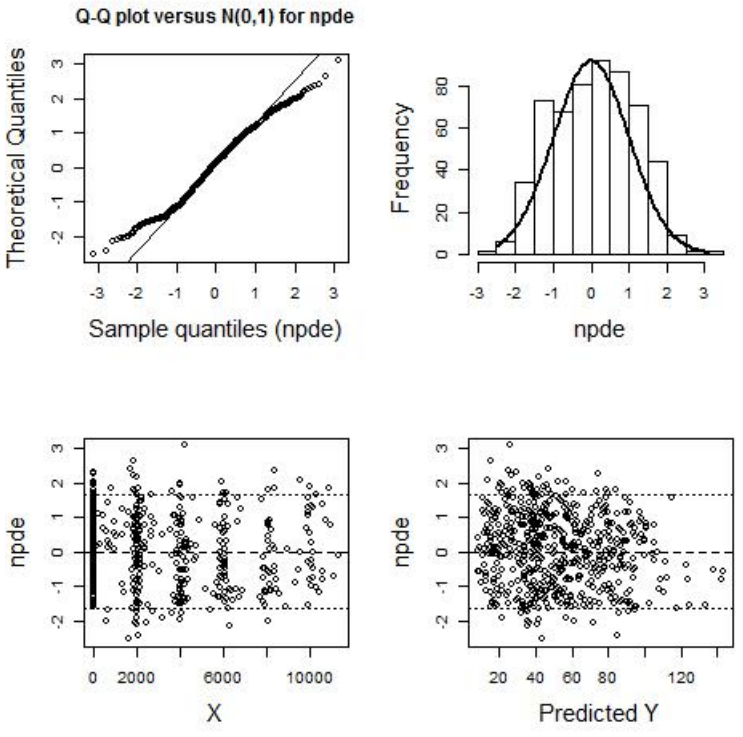


	OBJ	MIN?	Cov?	Corr?	SigDig	-LOG(NPDE)	Rank	N Parms	Crowding	Files	Plots	Gen	Ind	Density
9001	4795.6	Yes	No	No	3	2.27	1	8	0.7576			31	1	29
9002	5002.7	Yes	No	No	3.1	1.37	1	15	0.7576			31	2	6
9003	4746.7	Yes	Yes	Yes	3.4	3.3	1	6	0.0364			31	3	181
9004	4770.3	Yes	Yes	Yes	3.5	4.4	1	5	1.2157			31	4	106
9005	4735.9	Yes	Yes	Yes	3	4.13	1	10	0.1386			31	5	174
9006	4741.3	Yes	No	No	3.3	2.83	1	9	0.0928			31	6	106
9007	4738.5	No	No	No	1.1	3.22	1	12	0.1484			31	7	71
9008	4742.8	Yes	Yes	Yes	3.8	2.82	1	8	0.0364			31	8	191
9009	4743.5	Yes	Yes	Yes	3.5	3.02	1	7	0.0364			31	9	202
9010	4733.	No	No	No	2.5	2.95	1	16	0.7255			31	10	58
9011	4738.4	No	No	No	2.7	3.31	1	11	0.0574			31	11	139
9012	4741.6	Yes	Yes	Yes	3.7	3.21	2	9	0.069			31	12	185
9013	4739.5	Yes	No	No	3	3.06	2	12	0.0199			31	13	59
9014	4738.8	No	No	No	2.3	3.1	1	10	0.0577			31	14	175
9015	4745.3	No	No	No	1.2	2.48	2	11	0.0679			31	15	89
9016	4740.8	No	No	No	2.5	2.5	2	10	0.0007			31	16	122
9017	4741.6	Yes	Yes	Yes	3.2	3.65	2	9	0.0805			31	17	196
9018	4738.7	Yes	Yes	Yes	3.3	3.9	1	10	0.037			31	18	115
9019	4746.7	No	No	No	0.2	2.3	1	12	0.0574			31	19	97
9020	4747.	No	No	No	1.7	2.34	1	10	0.1481			31	20	144
9021	4742.9	Yes	Yes	Yes	3.3	2.77	1	8	0.0388			31	21	168
9022	4741.7	Yes	Yes	Yes	3	2.87	1	11	0.0367			31	22	110
9023	4742.3	Yes	Yes	Yes	3.4	2.89	1	8	0.0364			31	23	182
9024	4732.2	Yes	No	No	3.1	3.73	1	14	0.6667			31	24	87
9025	4740.5	No	No	No	2.8	2.71	1	9	0.0584			31	25	197
9026	4740.	No	No	No	2.3	3.43	1	9	0.0579			31	26	170
9027	4740.8	Yes	Yes	Yes	3.8	2.96	2	11	0.1847			31	27	94
9028	4741.6	No	No	No	2.9	3.01	1	8	0.0581			31	28	176
9029	4731.7	No	No	No	2.3	3.35	1	14	0.8147			31	29	118
9030	4737.7	No	No	No	0.7	3.15	1	13	0.1505			31	30	59
9031	4738.7	Yes	Yes	Yes	3.1	3.39	1	11	0.0364			31	31	103
9032	4742.3	Yes	Yes	Yes	3.1	2.87	1	8	0.0364			31	32	225
9033	4740.2	Yes	No	No	3.3	3.11	2	11	0.0048			31	33	44
9034	4740.2	Yes	No	No	3.2	2.8	1	11	0.0008			31	34	117
9035	4740.4	Yes	Yes	Yes	3.2	3.23	2	11	0.069			31	35	115
9036	4734.2	Yes	Yes	Yes	3.1	3.83	1	12	0.0364			31	36	158
9037	4741.7	Yes	Yes	Yes	3.1	2.49	1	9	0.0364			31	37	211
9038	4740.6	Yes	Yes	Yes	3.2	2.96	1	10	0.0374			31	38	149
9039	4739.7	Yes	Yes	Yes	3.4	3.67	1	9	0.0386			31	39	187
9040	4732.9	Yes	No	No	3.3	3.73	1	11	0.0012			31	40	119

Generation 31



Model 2, Failed Cov, Best NPDE = 1.37



Model 5, Best OBJ with Successful Cov, NPDE = 4.13

Generated Control Files for Models 2 and 5

```
TVCL3 = (THETA(1) +AGE*THETA(4)) *EXP(WT*THETA(5))
TVCL2 = TVCL3 +SEX*THETA(6)
TVCL1 = TVCL2  *(1+AI*THETA(7))
TVCL = TVCL1 *(1+CI*THETA(8))
CL   = TVCL +ETA(1)
TVV  = THETA(2)  +SEX*THETA(9)
TV    = TVV *(1+ETA(2))
TVKA = THETA(3)
KA   = TVKA +ETA(3)
V = TV
S2   = V
```

Model 2, OBJ = 5003, Failed Cov, Best NPDE = 1.37

```
TVCL3 = THETA(1)*(1+AGE*THETA(4)) *(1+WT*THETA(5))
TVCL2 = TVCL3 +SEX*THETA(6)
TVCL1 = TVCL2
TVCL = TVCL1  +CII*THETA(7)
CL   = TVCL *EXP(ETA(1))
TVV  = THETA(2)  +SEX*THETA(8)
TV    = TVV *EXP(ETA(2))
TVKA = THETA(3)
KA   = TVKA
V = TV
S2   = V
```

Model 5, Best OBJ with Successful Cov, NPDE = 4.13

Context:

- To do this, you need to be able to create hypotheses in large batches
- So, this really isn't very applicable to highly exploratory modeling, where hypotheses are often generated one at a time.
- This is best suited to the fairly routine modeling work (things that usually don't get presented in meetings like this)

An Odyssey through space (actual space or model search space) (Spoiler alert)

Next steps:

- Currently a stable Windows application for SOHGA - ready to share
- MOGA not yet ready to share
- Web application for (Monolix??) (SOHGA/MOGA?)


"...this conversation can serve no purpose anymore Goodbye."
HAL in
"2001 - A Space Odyssey"
© CREATIVE-WALLPAPERS.COM