DESIGN OF EXPERIMENTS FOR NONLINEAR KNOWLEDGE-BASED MODELS CALCULATED WITH NEURO PEX

Minimizing the number of trials in the industry and in laboratories

Terms of the problem:

Let's find the optimal points of experiments to estimate with a high accuracy the model parameters of Porphyridium's cellular growth model described with the following information:

2 input variables with their variation domain:

X_1 in [0..500] X_2 in [0..7]

4 parameters with an initial estimation of their possible values:

a = 0,0451   d = 20,96   b = 32,25   c = 0,17

Algebraic form of the postulated model:

$$Y = a\frac{X_1}{b+X_1}\frac{X_2}{c+X_2}\frac{d}{d+X_2}$$

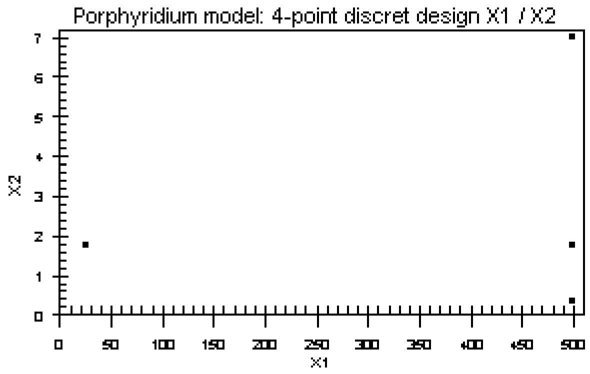
Solution :

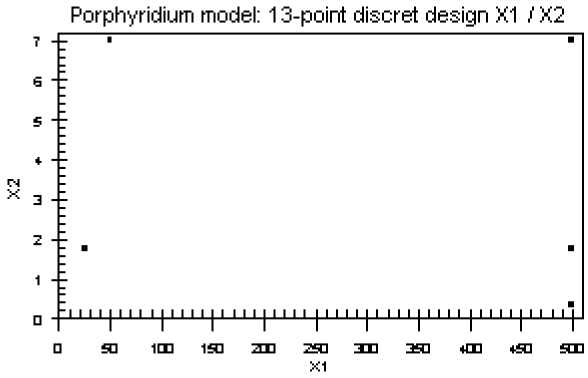
4- and 13-points D-optimal designs of experiments calculated with NEURO PEX are:

4-points designs of experiments			13-points designs of experiments	
X_1	X_2		X_1	X_2
25	1.75	1	25	1.75
500	0.35	2	25	1.75
500	1.75	3	25	1.75
500	7	4	50	7
		5	500	0.35
		6	500	0.35
		7	500	0.35
		8	500	1.75
		9	500	1.75
		10	500	1.75
		11	500	7
		12	500	7
		13	500	7

NETRAL - S.A. au capital de 213 429 Euros - R.C.S. Nanterre B 394 912 737 (94B02423)
14 rue Verdi, 92130 Issy-Les-Moulineaux, France - Tél : 01 46 38 75 12 - Fax : 01 46 38 75 94 - www.netral.com

In the 2-dimension experimental domain [X_1 ; X_2], the representations of the D-optimaux points are (on the left: 4-point design; on the right: 13-point design):



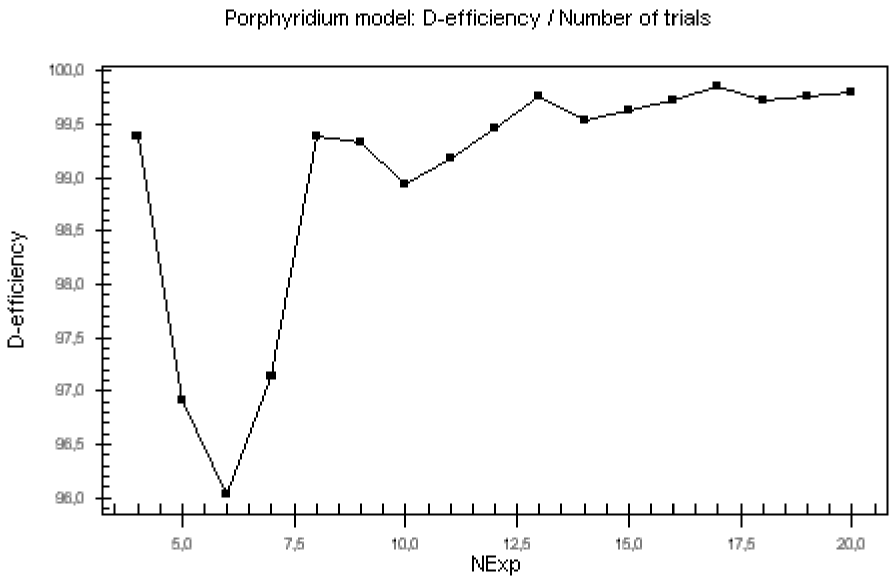


Remark that the software recommands to repeat points rather than to pave regularly the experimental domain. The result of it is an empty space, which does not spoil the model's quality. This phenomenon, already visible in a two-dimension space as the one presented above, becomes spectacularin spaces of greater dimension, i.e. in problems with three inputs or more.

The [50 ; 7] point appears from 13 experiments. This is due to probability of appearance in the D-optimal design, for a network of the exêrimental domain chosen by the user. NEURO PEX displays appearance probability table, here, for a 21x21 = 441-point network:

Treatment	X_1	X_2	Probability
1	500	0,35	0,246
2	500	1,75	0,243
3	500	7	0,229
4	25	1,75	0,211
5	50	7	0,062
6	50	0,35	0,009
7	25	1,4	6,47 E-6
8	0	0	0
9	25	0,35	0
10	50	0,70	0
…	…	…	…
…	..	…	…
440	500	6,30	0
441	500	6,65	0

Choosing 4 or 13 points is justified by the D-efficiency criterion. This criterion indicates how much information is conveyed by a trial to identified a model. Here, 4- or 8-trial designs conveys more information than 5-, 6- or 7-trial points. A 13-trial design will be better than 12- or 14-trial designs. Considering the available experimental budget, 4-, 8- or 13-trial designs will be chosen.



Conclusion and remarks:

After carrying out the 4-, 8- or 13-trial designs proposed above, and getting the associated Y responses, you should process to a least squares regression and to a new appraisal of the model parameters. Because of its construction, this new appraisal will be closer from the parameters real value. We will not detail this part which does not belong to the design of experiments, strictly speaking.

For a nonlinear model, the calculation of the design of experimentsso as to identify the model's parametersis possible only if a reasonnable appraisal of the parameter's value is already available. This well-known ambiguity is proper to nonlinear models. For an extremely precise identification, a sequential reasonning can be thought of, that would alternate experimenting and identifying parameters. The examples that have been treated by NETRAL indicate convergences in 1 to 3 iterations of the D-optimality algorithm, depending on the initial quality of the parameters appraisal, the extent of the measure noise and the low or high nonlinearity of the studied problem. For highly nonlinear models, X-optimality algorithms are to be prefered, even if the calculation time is longer.

EXAMPLE OF CALCULATED DESIGNS FOR A NEURAL NETWORK

See Netral's presentation in this page:

http://www.netral.com/lasociete/2004/2004-1201-conference-chimiometrie-si.pdf