

Figure S1. The main procedure of GA.

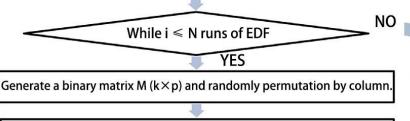
	Wavelength p, p = 1,2,···,10										
	1	1	1	1	1	1	1	1	1	1	
Sampling run k, k = 1,2,···,6	1	1	1	1	1	1	1	1	1	1	
	1	1	1	1	1	1	1	1	1	1	
	0	0	0	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	
The mean number of sampled							Permutation randomly				
	The m	ean nu	imber	of sam	piea	l Pe	rmutat	ion rai	ndomly	/	
	variab				. (rmutat each co		ndomly	/	
4 🔷					. (ndomly	0	
4 💠	variab		all runs	, Alpha	a = 5		each co	olumn			
	variab 0	les of a	all runs	, Alpha	a = 5	in 0	each co	olumn 1	1	0	
6	variab 0	les of a	o 0	0 1	a = 5 0 1	in 0	0	olumn 1 0	1	0 0	
6 4	variab 0 1 0	1 1 0	0 0	0 1 1	a = 5 0 1 0	1 0 0	0 1 0	1 0 1	1 1 0	0 0 0	

Each column has the same number of ones, indicating that each variable has the same chance of being sampled.

Figure S2. An example of the BMS approach.

LIBS dataset (p variables, n samples) (6144 \times 240 in this study)

Set the parameters of VCPA, including the number of sampling runs, k; the EDF runs, N; the mean number of sampled variables of all runs, Alpha; the ratio of best models of k sub-models, Ratio_better.



Calculate the RMSECV of each subset in each row of M.

Sort all the k RMSECV. Compute the frequency of variables from the best ratio subsets.

Compute the left number of variables using EDF. Retain the variables based on the frequency.

After N runs, there are L variables left for the further optimization



Figure S3. Flow chart of VCPA.