Supplementary material for the paper "The mRMR variable selection method: a comparative study for functional data"

José R. Berrendero, Antonio Cuevas, José L. Torrecilla
Departamento de Matemáticas
Universidad Autónoma de Madrid, Spain

1 List of models used in the simulation study and a few example outputs

Our simulation study consists of 400 experiments based on 100 different underlying models. The optimal classification rule in each case depends only on a finite number of variables. Models differ in complexity and number of relevant variables. The processes involved are chosen among the following: first, the **standard Brownian Motion**, B. Second, BT denotes a **Brownian Motion with a trend** $m(t)$, i.e., $BT(t) = B(t)+m(t)$; we have considered several choices for $m(t)$, a linear trend, $m(t) = ct$, a linear trend with random slope, i.e., $m(t) = \theta t$, where θ is a Gaussian r.v., and different members of two parametric families: the *peak* functions $\Phi_{m,k}$ and the *hillside* functions, defined by

$$\Phi_{m,k} = \int_0^t \varphi_{m,k}(s)ds \quad , \quad \text{hillside}_{t_0,b}(t) = b(t-t_0)\mathbb{I}_{[t_0,\infty)},$$

where, $\varphi_{m,k}(t) = \sqrt{2^{m-1}}\left[\mathbb{I}_{\left(\frac{2k-2}{2^m},\frac{2k-1}{2^m}\right)} - \mathbb{I}_{\left(\frac{2k-1}{2^m},\frac{2k}{2^m}\right)}\right]$ for $m \in \mathbb{N}$, $1 \le k \le 2^{m-1}$. Third, the **Brownian Bridge**: $BB(t) = B(t)-tB(1)$. Our fourth class of Gaussian processes is the **Ornstein–Uhlenbeck process**, with zero mean (OU) or different mean functions $m(t)$ (OUt). Finally some "smooth" processes have been also include. They are obtained by convolving Brownian trajectories with Gaussian kernels. We have considered two levels of smoothing denoted by sB and ssB.

In the following list of models, μ_i denotes de distribution of $X|Y = i$ and *variables* is the set of relevant variables in each Gaussian or Mixture case. We call them "relevant" in the sense that the optimal classification rule depends only on these variables. In the list below the variables written in boldface are "especially relevant" regarding their influence in the optimal classifier.

1. Gaussian models considered:

1. **G1** : $\begin{cases}\mu_0: & B(t)\\ \mu_1: & B(t)+\theta t \quad ,\theta\sim N(0,3)\end{cases}$

   $variables = \{X_{100}\}$.

2. **G1b** : $\begin{cases}\mu_0: & B(t)\\ \mu_1: & B(t)+\theta t \quad ,\theta\sim N(0,5)\end{cases}$

   $variables = \{X_{100}\}$.

3. **G2** : $\begin{cases}\mu_0: & B(t)+t\\ \mu_1: & B(t)\end{cases}$

   $variables = \{X_{100}\}$.

4. **G2b** : $\begin{cases}\mu_0: & B(t)+3t\\ \mu_1: & B(t)\end{cases}$

   $variables = \{X_{100}\}$.

5. $\mathbf{G3}: \begin{cases} \mu_0: & BB(t) \\ \mu_1: & B(t) \end{cases}$

$variables = \{X_{100}\}$.

6. $\mathbf{G4}: \begin{cases} \mu_0: & B(t) + hillside_{0.5,4}(t) \\ \mu_1: & B(t) \end{cases}$

$variables = \{X_{47}, \boldsymbol{X}_{100}\}$.

7. $\mathbf{G5}: \begin{cases} \mu_0: & B(t) + 3\Phi_{1,1}(t) \\ \mu_1: & B(t) \end{cases}$

$variables = \{X_1, \boldsymbol{X}_{48}, X_{100}\}$.

8. $\mathbf{G6}: \begin{cases} \mu_0: & B(t) + 5\Phi_{2,2}(t) \\ \mu_1: & B(t) \end{cases}$

$variables = \{X_{48}, \boldsymbol{X}_{75}, X_{100}\}$.

9. $\mathbf{G7}: \begin{cases} \mu_0: & B(t) + 5\Phi_{3,2}(t) + 5\Phi_{3,4}(t) \\ \mu_1: & B(t) \end{cases}$

$variables = \{X_{22}, \boldsymbol{X}_{35}, X_{49}, X_{74}, \boldsymbol{X}_{88}, X_{100}\}$.

10. $\mathbf{G8}: \begin{cases} \mu_0: & B(t) + 3\Phi_{2,1.25}(t) + 3\Phi_{2,2}(t) \\ \mu_1: & B(t) \end{cases}$

$variables = \{X_9, \boldsymbol{X}_{35}, X_{48}, X_{62}, \boldsymbol{X}_{75}, X_{100}\}$.

2. LOGISTIC-TYPE MODELS CONSIDERED: they are all defined according standard (ii) (see Sec. 4 in the main paper). The process $X = X(t)$ follows one of the distributions mentioned above and $Y = \text{Binom}(1, \eta(X))$ with $\eta(x) = (1 + e^{-\psi(x(t_1),\cdots,x(t_k))})^{-1}$, a function of the relevant variables $x(t_1), \cdots, x(t_k)$.

L1: $\psi(X) = 10X_{65}$.

L2: $\psi(X) = 10X_{30} + 10X_{70}$.

L3: $\psi(X) = 10X_{30} - 10X_{70}$.

L4: $\psi(X) = 20X_{30} + 50X_{50}20X_{80}$.

L5: $\psi(X) = 20X_{30} - 50X_{50} + 20X_{80}$.

L6: $\psi(X) = 10X_{10} + 30X_{40} + 10X_{72} + 10X_{80} + 20X_{95}$.

L7: $\psi(X) = \sum_{i=1}^{10} 10X_{10i}$.

L8: $\psi(X) = 20X_{30}^2 + 10X_{50}^4 + 50X_{80}^3$.

L9: $\psi(X) = 10X_{10} + 10|X_{50}| + 0X_{30}^2X_{85}$.

L10: $\psi(X) = 20X_{33} + 20|X_{68}|$.

L11: $\psi(X) = \frac{20}{X_{35}} + \frac{30}{X_{77}}$.

L12: $\psi(X) = \log X_{35} + \log X_{77}$.

L13: $\psi(X) = 40X_{20} + 30X_{28} + 20X_{62} + 10X_{67}$.

L14: $\psi(X) = 40X_{20} + 30X_{28} - 20X_{62} - 10X_{67}$.

L15: $\psi(X) = 40X_{20} - 30X_{28} + 20X_{62} - 10X_{67}$.

Some variations of these models have been also considered:

L3b: $\psi(X) = 30X_{30} - 20X_{70}$.

L4b: $\psi(X) = 30X_{30} + 20X_{50} + 10X_{80}$.

L5b: $\psi(X) = 10X_{30} - 10X_{50} + 10X_{80}$.

L6b: $\psi(X) = 20X_{10} + 20X_{40} + 20X_{72} + 20X_{80} + 20X_{95}$.

L8b: $\psi(X) = 10X_{30}^2 + 10X_{50}^4 + 10X_{80}^3$.

3. Mixture-type models: they are obtained by combining (via mixtures) in several ways the above mentioned Gaussian distributions assumed for $X|Y=0$ and $X|Y=1$. These models are denoted M1, ..., M10 in the output tables.

1. $\mathbf{M1} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3t, & 1/2 \\ B(t)-2t, & 1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_{100}\}$.

2. $\mathbf{M2} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{2,2}(t), & 1/2 \\ B(t)+5\Phi_{3,2}(t), & 1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_{22}, \boldsymbol{X}_{35}, X_{48}, \boldsymbol{X}_{75}, X_{100}\}$.

3. $\mathbf{M3} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{2,2}(t), & 1/10 \\ B(t)+5\Phi_{3,2}(t), & 9/10 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_{22}, \boldsymbol{X}_{35}, X_{48}, \boldsymbol{X}_{75}, X_{100}\}$.

4. $\mathbf{M4} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{2,2}(t), & 1/2 \\ B(t)+5\Phi_{3,3}(t), & 1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_{48}, \boldsymbol{X}_{62}, \boldsymbol{X}_{75}, X_{100}\}$.

5. $\mathbf{M5} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{2,1}(t) & ,1/3 \\ B(t)+3\Phi_{2,2}(t), & 1/3 \\ B(t)+5\Phi_{3,2}(t), & 1/3 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_1, \boldsymbol{X}_{22}, \boldsymbol{X}_{35}, X_{48}, \boldsymbol{X}_{75}, X_{100}\}$.

6. $\mathbf{M6} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{2,1}(t) & ,1/2 \\ B(t)+3t & ,1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_1, \boldsymbol{X}_{22}, X_{49}, \boldsymbol{X}_{100}\}$.

7. $\mathbf{M7} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{1,1}(t) & ,1/2 \\ BB(t) & ,1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_1, \boldsymbol{X}_{48}, \boldsymbol{X}_{100}\}$.

8. $\mathbf{M8} : \begin{cases} \mu_0 : & \begin{cases} B(t)+\theta t, \ \ \theta \sim N(0,5) & ,1/2 \\ B(t)+hillside_{0.5,5}(t) & ,1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = \{X_{47}, \boldsymbol{X}_{100}\}$.

9. $\mathbf{M9} : \begin{cases} \mu_0 : & \begin{cases} B(t)+\theta t, \ \ \theta \sim N(0,5) & ,1/2 \\ BB(t) & ,1/2 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

   $variables = X_{100}$.

10. $\mathbf{M10} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{1,1}(t) & ,1/3 \\ B(t)-3t & ,1/3 \\ BB(t) & ,1/3 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

    $variables = \{X_1, \boldsymbol{X}_{48}, \boldsymbol{X}_{100}\}$.

11. $\mathbf{M11} : \begin{cases} \mu_0 : & \begin{cases} B(t)+3\Phi_{1,1}(t) & ,1/4 \\ B(t)-3t & ,1/4 \\ B(t)+hillside_{0.5,5}(t) & ,1/4 \\ BB(t) & ,1/4 \end{cases} \\ \mu_1 : & B(t) \end{cases}$

    $variables = \{X_1, \boldsymbol{X}_{48}, \boldsymbol{X}_{100}\}$.

Finally, the full list of models involved is as follows:

1. L1 OU
2. L1 OUt
3. L1 B
4. L1 sB
5. L1 ssB
6. L2 OU
7. L2 OUt
8. L2 B
9. L2 sB
10. L2 ssB
11. L3 OU
12. L3b OU
13. L3 OUt
14. L3b OUt
15. L3 B
16. L3b B
17. L3 sB
18. L3 ssB
19. L4 OU
20. L4b OU
21. L4 OUt
22. L4b OUt
23. L4 B
24. L4 sB
25. L4 ssB
26. L5 OU
27. L5b OU
28. L5 OUt
29. L5 B
30. L5 sB
31. L5 ssB
32. L6 OU
33. L6b OU
34. L6 OUt
35. L6b OUt
36. L6 B
37. L6 sB
38. L6 ssB
39. L7 OU
40. L7b OU
41. L7 OUt
42. L7b OUt
43. L7 B
44. L7 sB
45. L7 ssB
46. L8 B
47. L8 sB
48. L8 ssB
49. L8b OU
50. L9 B
51. L9 sB
52. L9 ssB
53. L10 OU
54. L10 B
55. L10 sB
56. L10 ssB
57. L11 OU
58. L11 OUt
59. L11 B
60. L11 sB
61. L11 ssB
62. L12 OU
63. L12 OUt
64. L12 B
65. L12 sB
66. L12 ssB
67. L13 OU
68. L13 OUt
69. L13 B
70. L13 sB
71. L13 ssB
72. L14 OU
73. L14 OUt
74. L14 B
75. L14 sB
76. L15 OU
77. L15 OUt
78. L15 B
79. L15 sB
80. G1
81. G1b
82. G2
83. G2b
84. G3
85. G4
86. G5
87. G6
88. G7
89. G8
90. M1
91. M2
92. M3
93. M4
94. M5
95. M6
96. M7
97. M8
98. M9
99. M10
100. M11

We next provide a few simulation results. See `www.uam.es/antonio.cuevas/exp/mRMR-outputs.xlsx` for the full simulation outputs.

NB accuracy outputs

Model	MID	FCD	RD	VD	CD	Base
L7_OU	$\boxed{90.73}$	87.01	90.41	87.72	90.50	$\boxed{93.35}$
L1_OUt	72.54	$\boxed{74.92}$	74.10	$\boxed{74.56}$	74.26	73.33
L14_B	75.89	$\boxed{77.31}$	$\boxed{77.17}$	76.72	76.75	75.33
L9_sB	84.73	$\boxed{86.01}$	$\boxed{85.84}$	85.40	85.78	82.27
G1b	79.74	75.94	$\boxed{80.73}$	$\boxed{81.66}$	77.67	79.49
G3	76.22	64.54	$\boxed{78.84}$	$\boxed{78.78}$	72.67	71.24
G6	83.13	83.50	$\boxed{84.38}$	83.96	$\boxed{84.28}$	74.90
M1	79.12	73.09	$\boxed{80.93}$	$\boxed{81.89}$	76.77	77.61
M4	64.73	$\boxed{68.04}$	$\boxed{67.91}$	67.48	67.89	61.36
M6	81.25	80.09	$\boxed{83.06}$	82.45	$\boxed{83.21}$	78.02

Table 10.- Average NB accuracy (proportion of correct classification) outputs, over 200 runs of the considered methods with sample size $n = 50$.

NB number of variables

Model	MID	FCD	RD	VD	CD	Base
L7_OU	12.1	14.1	$\boxed{10.1}$	11.5	$\boxed{11.1}$	100
L1_OUt	8.9	8.0	7.0	$\boxed{5.9}$	$\boxed{6.6}$	100
L14_B	7.9	8.0	$\boxed{6.0}$	7.5	$\boxed{5.9}$	100
L9_sB	$\boxed{4.9}$	7.5	$\boxed{4.7}$	4.9	5.3	100
G1b	6.9	13.4	$\boxed{4.8}$	$\boxed{1.5}$	9.0	100
G3	6.4	11.6	$\boxed{3.7}$	$\boxed{3.7}$	8.2	100
G6	3.4	$\boxed{1.9}$	3.5	3.1	$\boxed{2.7}$	100
M1	4.9	11.7	$\boxed{4.5}$	$\boxed{1.4}$	8.4	100
M4	6.5	5.0	5.1	$\boxed{3.4}$	$\boxed{4.6}$	100
M6	7.1	7.6	$\boxed{5.5}$	$\boxed{5.6}$	6.4	100

Table 11.- Average number of selected variables over 200 runs of the considered methods with sample size $n = 50$ using NB.

k-NN accuracy outputs

Model	MID	FCD	RD	VD	CD	Base
L7_OU	90.74	86.89	$\boxed{90.78}$	87.79	90.67	$\boxed{92.21}$
L1_OUt	75.83	$\boxed{77.34}$	76.77	$\boxed{77.22}$	77.11	75.81
L14_B	75.34	$\boxed{77.16}$	$\boxed{77.29}$	76.20	76.49	74.43
L9_sB	86.79	$\boxed{87.39}$	$\boxed{87.35}$	87.03	87.28	86.10
G1b	79.11	75.74	$\boxed{80.00}$	$\boxed{80.01}$	78.13	78.57
G3	73.39	61.97	$\boxed{77.28}$	$\boxed{77.03}$	68.20	65.26
G6	$\boxed{91.95}$	84.16	88.22	85.80	84.68	$\boxed{92.19}$
M1	81.63	75.47	$\boxed{82.79}$	$\boxed{83.00}$	80.59	80.72
M4	72.94	71.60	$\boxed{74.72}$	70.30	71.66	$\boxed{73.29}$
M6	83.08	79.70	$\boxed{84.13}$	$\boxed{84.26}$	84.02	80.99

Table 12.- Average k-NN accuracy (proportion of correct classification) outputs, over 200 runs of the considered methods with sample size $n = 50$.

k-NN number of variables

Model	MID	FCD	RD	VD	CD	Base
L7_OU	11.5	14.5	$\boxed{10.4}$	12.1	$\boxed{11.3}$	100
L1_OUt	9.0	7.9	6.9	$\boxed{6.5}$	$\boxed{6.8}$	100
L14_B	8.3	7.6	$\boxed{5.5}$	8.0	$\boxed{6.5}$	100
L9_sB	6.3	7.7	$\boxed{6.0}$	7.1	$\boxed{6.0}$	100
G1b	7.8	11.7	$\boxed{6.5}$	$\boxed{6.3}$	8.7	100
G3	5.1	11.2	$\boxed{2.5}$	$\boxed{2.9}$	7.8	100
G6	11.5	12.6	9.0	$\boxed{8.2}$	$\boxed{7.5}$	100
M1	7.8	11.4	$\boxed{6.3}$	$\boxed{4.9}$	8.6	100
M4	11.7	16.1	10.4	$\boxed{9.7}$	$\boxed{10.1}$	100
M6	9.9	9.6	$\boxed{7.5}$	8.7	$\boxed{7.6}$	100

Table 13.- Average number of selected variables over 200 runs of the considered methods with sample size $n = 50$ using k-NN.

LDA accuracy outputs

Model	MID	FCD	RD	VD	CD	Base
L7_OU	89.27	85.13	$\boxed{89.75}$	86.81	$\boxed{90.10}$	-
L1_OUt	72.05	$\boxed{73.74}$	73.48	$\boxed{73.96}$	73.66	-
L14_B	75.25	$\boxed{76.35}$	$\boxed{77.12}$	75.62	76.27	-
L9_sB	84.91	84.88	$\boxed{84.96}$	84.69	$\boxed{85.12}$	-
G1b	53.35	52.22	$\boxed{54.15}$	$\boxed{54.49}$	51.67	-
G3	52.26	50.91	$\boxed{53.53}$	$\boxed{53.51}$	51.06	-
G6	$\boxed{95.28}$	87.92	$\boxed{90.54}$	86.59	87.80	-
M1	54.44	53.64	$\boxed{54.88}$	$\boxed{54.68}$	53.70	-
M4	$\boxed{78.95}$	70.98	$\boxed{76.46}$	71.37	71.30	-
M6	79.92	79.53	$\boxed{80.80}$	80.57	$\boxed{80.70}$	-

Table 14.- Average LDA accuracy (proportion of correct classification) outputs, over 200 runs of the considered methods with sample size $n = 50$.

LDA number of variables

Model	MID	FCD	RD	VD	CD	Base
L7_OU	5.5	6.9	6.2	6.1	7.2	-
L1_OUt	5.6	4.6	4.6	4.4	4.9	-
L14_B	3.6	3.2	3.1	4.1	3.6	-
L9_sB	4.0	3.5	3.3	3.4	3.5	-
G1b	5.6	7.0	5.3	4.6	7.4	-
G3	7.1	8.9	5.0	5.0	8.9	-
G6	10.7	14.3	11.2	7.8	11.6	-
M1	5.5	7.2	5.3	5.7	6.8	-
M4	11.1	10.2	10.3	7.4	8.5	-
M6	6.4	4.5	5.2	5.2	4.8	-

Table 15.- Average number of selected variables over 200 runs of the considered methods with sample size $n = 50$ using LDA.

SVM accuracy outputs

Model	MID	FCD	RD	VD	CD	Base
L7_OU	90.73	87.01	90.41	87.72	90.50	93.35
L1_OUt	72.54	74.92	74.10	74.56	74.26	73.33
L14_B	75.89	77.31	77.17	76.72	76.75	75.33
L9_sB	84.73	86.01	85.84	85.40	85.78	82.27
G1b	79.74	75.94	80.73	81.66	77.67	79.49
G3	76.22	64.54	78.84	78.78	72.67	71.24
G6	83.13	83.50	84.38	83.96	84.28	74.90
M1	79.12	73.09	80.93	81.89	76.77	77.61
M4	64.73	68.04	67.91	67.48	67.89	61.36
M6	81.25	80.09	83.06	82.45	83.21	78.02

Table 16.- Average SVM accuracy (proportion of correct classification) outputs, over 200 runs of the considered methods with sample size $n = 50$.

SVM number of variables

Model	MID	FCD	RD	VD	CD	Base
L7_OU	12.1	14.1	10.1	11.5	11.1	100
L1_OUt	8.9	8.0	7.0	5.9	6.6	100
L14_B	7.9	8.0	6.0	7.5	5.9	100
L9_sB	4.9	7.5	4.7	4.9	5.3	100
G1b	6.9	13.4	4.8	1.5	9.0	100
G3	6.4	11.6	3.7	3.7	8.2	100
G6	3.4	1.9	3.5	3.1	2.7	100
M1	4.9	11.7	4.5	1.4	8.4	100
M4	6.5	5.0	5.1	3.4	4.6	100
M6	7.1	7.6	5.5	5.6	6.4	100

Table 17.- Average number of selected variables over 200 runs of the considered methods with sample size $n = 50$ using SVM.

2 Results with quotient criterion and other classifiers

The outputs included in the paper correspond to the difference criterion (in order to combine the relevance and redundancy measures). We provide here some additional results for the quotient criterion, instead of the difference one. Besides, Figure 2 in the paper was produced using only the outputs for the k-NN classfier. In this document we show the analogous displays for LDA, SVM and NB. Again, the entire simulation outputs can be downloaded from `www.uam.es/antonio.cuevas/exp/mRMR-outputs.xlsx`.

Tables 1a-9a below correspond and Tables 1-9 in the main paper with the difference criterion replaced with the quotient one. Figures 2a, 2b and 2c correspond to Figure 2 using the other classifiers.

Output (NB)	Sample size	MIQ	FCQ	RQ	VQ	CQ	Base
Avgerage accuracy	$n = 30$	78.10	78.76	79.53	79.58	79.12	77.28
	$n = 50$	79.59	79.73	80.86	80.81	80.26	78.29
	$n = 100$	80.62	80.54	81.82	81.75	81.16	78.84
	$n = 200$	81.24	81.03	82.48	82.35	81.77	79.13
Average dim. red	$n = 30$	8.9	8.6	7.2	7.0	7.9	100
	$n = 50$	8.3	8.1	6.7	6.7	7.4	100
	$n = 100$	7.7	7.2	6.1	6.1	6.9	100
	$n = 200$	7.1	6.7	5.7	5.9	6.5	100
Victories over Base	$n = 30$	60	65	72	76	68	-
	$n = 50$	67	61	79	78	68	-
	$n = 100$	71	64	88	86	79	-
	$n = 200$	75	68	92	91	84	-

Table 1a.- Performance outputs for the considered methods, using NB and the quotient criterion, with different sample sizes. Each output is the result of the 100 different models for each sample size.

Output (k-NN)	Sample size	MIQ	FCQ	RQ	VQ	CQ	Base
Avgerage accuracy	$n = 30$	80.02	79.65	80.85	80.82	80.09	78.98
	$n = 50$	81.32	80.40	81.72	81.67	80.87	80.34
	$n = 100$	82.84	81.34	82.73	82.65	81.83	81.99
	$n = 200$	84.06	82.09	83.56	83.49	82.65	83.38
Average dim. red	$n = 30$	9.3	9.5	7.4	7.6	8.1	100
	$n = 50$	9.6	9.6	7.6	7.8	8.3	100
	$n = 100$	9.9	9.9	8.0	8.2	8.6	100
	$n = 200$	10.1	10.1	8.3	8.5	9.0	100
Victories over Base	$n = 30$	71	58	76	74	67	-
	$n = 50$	67	53	73	72	64	-
	$n = 100$	71	49	64	64	55	-
	$n = 200$	64	42	62	64	54	-

Table 2a.- Performance outputs for the considered methods, using k-NN and the quotient criterion, with different sample sizes. Each output is the result of the 100 different models for each sample size.

Output (LDA)	Sample size	MIQ	FCQ	RQ	VQ	CQ	Base
Avgerage accuracy	$n = 30$	$\boxed{78.67}$	77.58	$\boxed{78.64}$	$\boxed{78.64}$	78.15	60.80
	$n = 50$	$\boxed{80.20}$	78.53	$\boxed{79.58}$	79.52	79.11	58.75
	$n = 100$	$\boxed{81.74}$	79.62	$\boxed{80.65}$	80.52	80.25	53.11
	$n = 200$	$\boxed{82.90}$	80.47	$\boxed{81.53}$	81.35	81.10	73.25
Average dim. red	$n = 30$	5.8	$\boxed{4.7}$	$\boxed{4.7}$	$\boxed{4.7}$	5.1	100
	$n = 50$	6.9	5.7	$\boxed{5.6}$	$\boxed{5.6}$	6.0	100
	$n = 100$	8.3	7.1	$\boxed{6.9}$	$\boxed{7.0}$	7.3	100
	$n = 200$	9.5	8.3	$\boxed{8.0}$	$\boxed{8.1}$	8.4	100

Table 3a.- Performance outputs for the considered methods, using LDA and the quotient criterion, with different sample sizes. Each output is the result of the 100 different models for each sample size.

Output (SVM)	Sample size	MIQ	FCQ	RQ	VQ	CQ	Base
Avgerage accuracy	$n = 30$	$\boxed{81.62}$	79.81	80.69	80.65	80.27	$\boxed{81.91}$
	$n = 50$	$\boxed{82.69}$	80.42	81.43	81.35	80.96	$\boxed{82.99}$
	$n = 100$	$\boxed{83.80}$	81.21	82.20	82.12	81.76	$\boxed{84.11}$
	$n = 200$	$\boxed{84.61}$	81.79	82.90	82.76	82.42	$\boxed{84.91}$
Average dim. red	$n = 30$	10.6	10.3	$\boxed{9.1}$	$\boxed{9.2}$	9.5	100
	$n = 50$	10.7	10.4	$\boxed{9.3}$	$\boxed{9.4}$	9.7	100
	$n = 100$	11.1	10.5	$\boxed{9.5}$	$\boxed{9.7}$	9.9	100
	$n = 200$	11.4	10.7	$\boxed{9.7}$	$\boxed{9.9}$	10.0	100
Victories over Base	$n = 30$	32	37	$\boxed{49}$	$\boxed{47}$	42	-
	$n = 50$	35	34	$\boxed{51}$	$\boxed{52}$	44	-
	$n = 100$	35	33	$\boxed{51}$	$\boxed{50}$	48	-
	$n = 200$	33	31	$\boxed{52}$	$\boxed{51}$	48	-

Table 4a.- Performance outputs for the considered methods, using SVM and the quotient criterion, with different sample sizes. Each output is the result of the 100 different models for each sample size.

Ranking criterion (NB)	Sample size	MIQ	FCQ	RQ	VQ	CQ
Relative	$n = 30$	2.27	5.67	$\boxed{8.69}$	$\boxed{8.63}$	7.65
	$n = 50$	2.69	4.94	$\boxed{9.09}$	$\boxed{8.70}$	7.51
	$n = 100$	2.75	4.75	$\boxed{9.21}$	$\boxed{8.80}$	7.44
	$n = 200$	2.71	4.57	$\boxed{8.87}$	$\boxed{8.31}$	7.41
Positional	$n = 30$	6.78	7.83	$\boxed{8.53}$	$\boxed{8.50}$	8.39
	$n = 50$	6.79	7.57	$\boxed{8.93}$	$\boxed{8.51}$	8.22
	$n = 100$	6.80	7.47	$\boxed{9.01}$	$\boxed{8.58}$	8.14
	$n = 200$	6.84	7.56	$\boxed{8.92}$	$\boxed{8.43}$	8.25
F1	$n = 30$	12.25	15.85	$\boxed{17.39}$	$\boxed{17.58}$	17.01
	$n = 50$	12.24	14.90	$\boxed{19.19}$	$\boxed{17.37}$	16.35
	$n = 100$	12.35	14.60	$\boxed{19.67}$	$\boxed{17.38}$	16
	$n = 200$	12.49	14.92	$\boxed{19.28}$	$\boxed{16.86}$	16.45

Table 5a.- Global scores of the considered (quotient-based) methods using three different ranking criteria with the NB classifier.

Ranking criterion (k-NN)	Sample size	MIQ	FCQ	RQ	VQ	CQ
Relative	$n = 30$	3.70	3.97	$\boxed{8.51}$	$\boxed{8.09}$	6.03
	$n = 50$	4.39	3.72	$\boxed{7.84}$	$\boxed{7.59}$	5.61
	$n = 100$	4.93	3.46	$\boxed{7.24}$	$\boxed{6.91}$	5.15
	$n = 200$	5.52	3.00	$\boxed{6.72}$	$\boxed{6.52}$	5.00
Positional	$n = 30$	7.31	7.29	$\boxed{9.06}$	$\boxed{8.64}$	7.70
	$n = 50$	7.55	7.30	$\boxed{8.90}$	$\boxed{8.67}$	7.58
	$n = 100$	7.75	7.37	$\boxed{8.82}$	$\boxed{8.62}$	7.49
	$n = 200$	7.96	7.43	$\boxed{8.56}$	$\boxed{8.45}$	7.60
F1	$n = 30$	14.32	13.96	$\boxed{19.66}$	$\boxed{17.80}$	14.26
	$n = 50$	15.27	14.06	$\boxed{18.78}$	$\boxed{17.94}$	13.95
	$n = 100$	16.02	14.19	$\boxed{18.44}$	$\boxed{17.90}$	13.62
	$n = 200$	16.84	14.09	$\boxed{17.48}$	$\boxed{17.57}$	14.02

Table 6a.- Global scores of the considered (quotient-based) methods using three different ranking criteria with the k-NN classifier.

Ranking criterion (LDA)	Sample size	MIQ	FCQ	RQ	VQ	CQ
Relative	$n = 30$	3.94	3.14	$\boxed{7.48}$	$\boxed{7.38}$	5.38
	$n = 50$	4.26	2.86	$\boxed{6.97}$	$\boxed{6.66}$	5.19
	$n = 100$	4.76	2.60	$\boxed{6.98}$	$\boxed{6.43}$	5.33
	$n = 200$	5.27	2.36	$\boxed{6.78}$	$\boxed{6.13}$	5.23
Positional	$n = 30$	7.49	6.99	$\boxed{9.01}$	$\boxed{8.96}$	7.55
	$n = 50$	7.64	7.12	$\boxed{8.90}$	$\boxed{8.64}$	7.70
	$n = 100$	7.72	7.13	$\boxed{8.89}$	$\boxed{8.52}$	7.74
	$n = 200$	7.80	7.23	$\boxed{8.79}$	$\boxed{8.36}$	7.82
F1	$n = 30$	15.05	12.67	$\boxed{19.11}$	$\boxed{19.32}$	13.85
	$n = 50$	15.63	12.95	$\boxed{18.91}$	$\boxed{18.07}$	14.44
	$n = 100$	15.80	13.04	$\boxed{18.83}$	$\boxed{17.72}$	14.61
	$n = 200$	16.25	13.29	$\boxed{18.62}$	$\boxed{17.04}$	14.80

Table 7a.- Global scores of the considered (quotient-based) methods using three different ranking criteria with LDA.

Ranking criterion (SVM)	Sample size	MIQ	FCQ	RQ	VQ	CQ
Relative	$n = 30$	6.02	2.85	$\boxed{6.58}$	$\boxed{6.28}$	4.42
	$n = 50$	5.99	2.72	$\boxed{6.70}$	$\boxed{6.15}$	4.73
	$n = 100$	$\boxed{6.14}$	2.61	$\boxed{6.43}$	5.99	4.66
	$n = 200$	$\boxed{6.42}$	2.30	$\boxed{6.29}$	5.75	4.68
Positional	$n = 30$	8.26	7.20	$\boxed{8.74}$	$\boxed{8.46}$	7.34
	$n = 50$	8.16	7.19	$\boxed{8.80}$	$\boxed{8.43}$	7.42
	$n = 100$	8.26	7.31	$\boxed{8.66}$	$\boxed{8.32}$	7.48
	$n = 200$	$\boxed{8.28}$	7.36	$\boxed{8.61}$	8.22	7.56
F1	$n = 30$	$\boxed{17.90}$	13.78	$\boxed{17.99}$	17.17	13.16
	$n = 50$	$\boxed{17.58}$	13.51	$\boxed{18.21}$	17.28	13.42
	$n = 100$	$\boxed{17.97}$	13.86	$\boxed{17.71}$	16.95	13.59
	$n = 200$	$\boxed{17.89}$	13.85	$\boxed{17.70}$	16.58	14.06

Table 8a.- Global scores of the considered (quotient-based) methods using three different ranking criteria with the linear SVM.

NB outputs

Output	Data	MIQ	FCQ	RQ	VQ	CQ	Base
Classification accuracy	Growth	$\boxed{88.17}$	$\boxed{87.10}$	86.02	86.02	87.10	84.95
	Tecator	96.28	97.67	$\boxed{99.53}$	$\boxed{99.53}$	98.14	97.21
	Phoneme	73.08	80.20	$\boxed{80.38}$	80.15	$\boxed{80.32}$	74.08
Number of variables	Growth	1.5	$\boxed{1.1}$	$\boxed{1.1}$	$\boxed{1.1}$	$\boxed{1.1}$	31
	Tecator	4.8	5.0	$\boxed{1.0}$	$\boxed{1.0}$	4.4	100
	Phoneme	14.5	$\boxed{10.6}$	16.7	16.8	$\boxed{14.1}$	256

k-NN outputs

Output	Data	MIQ	FCQ	RQ	VQ	CQ	Base
Classification accuracy	Growth	$\boxed{95.70}$	83.87	83.87	83.87	83.87	$\boxed{96.77}$
	Tecator	96.74	99.07	$\boxed{99.53}$	$\boxed{99.53}$	98.60	98.60
	Phoneme	75.53	$\boxed{81.42}$	79.79	80.38	$\boxed{80.61}$	78.80
Number of variables	Growth	3.9	$\boxed{1.0}$	$\boxed{1.0}$	$\boxed{1.0}$	$\boxed{1.0}$	31
	Tecator	4.0	3.0	$\boxed{1.0}$	$\boxed{1.0}$	4.3	100
	Phoneme	18.4	$\boxed{12.1}$	12.3	15.2	$\boxed{6.7}$	256

LDA outputs

Output	Data	MIQ	FCQ	RQ	VQ	CQ	Base
Classification accuracy	Growth	$\boxed{95.70}$	$\boxed{91.40}$	$\boxed{91.40}$	$\boxed{91.40}$	$\boxed{91.40}$	-
	Tecator	$\boxed{94.88}$	94.42	$\boxed{94.88}$	94.42	$\boxed{95.35}$	-
	Phoneme	74.55	78.88	79.10	$\boxed{79.63}$	$\boxed{80.26}$	-
Number of variables	Growth	$\boxed{3.7}$	5.0	$\boxed{4.9}$	$\boxed{4.9}$	5.0	-
	Tecator	6.1	8.4	4.1	$\boxed{2.2}$	$\boxed{3.1}$	-
	Phoneme	19.0	$\boxed{8.9}$	10.0	$\boxed{9.0}$	9.2	-

SVM outputs

Output	Data	MIQ	FCQ	RQ	VQ	CQ	Base
Classification accuracy	Growth	$\boxed{94.62}$	87.10	87.10	87.10	86.02	$\boxed{95.70}$
	Tecator	98.14	$\boxed{99.07}$	$\boxed{99.53}$	$\boxed{99.07}$	98.60	$\boxed{99.07}$
	Phoneme	75.30	$\boxed{80.71}$	80.67	80.37	80.33	$\boxed{80.96}$
Number of variables	Growth	$\boxed{3.5}$	5.0	$\boxed{4.9}$	$\boxed{4.9}$	5.0	31
	Tecator	6.7	2.1	$\boxed{1.0}$	$\boxed{1.0}$	4.1	100
	Phoneme	19.3	$\boxed{10.1}$	11.3	$\boxed{10.8}$	12.2	256

Table 9a.- Performances of the different (quotient-based) mRMR methods in three real data sets. From top to bottom tables stand for Naive Bayes, k-NN, LDA and linear SVM outputs respectively.

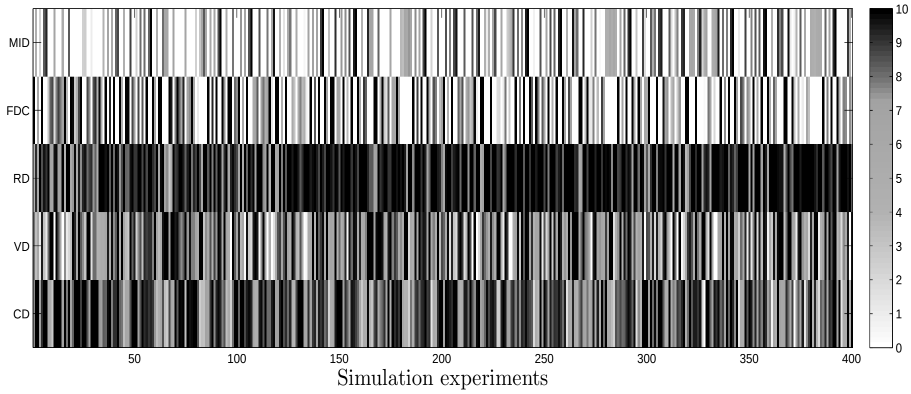

Figure 2a.- Cromatic version of the global relative ranking table taking into account the 400 considered experiments (columns) and the Naive Bayes classifier.

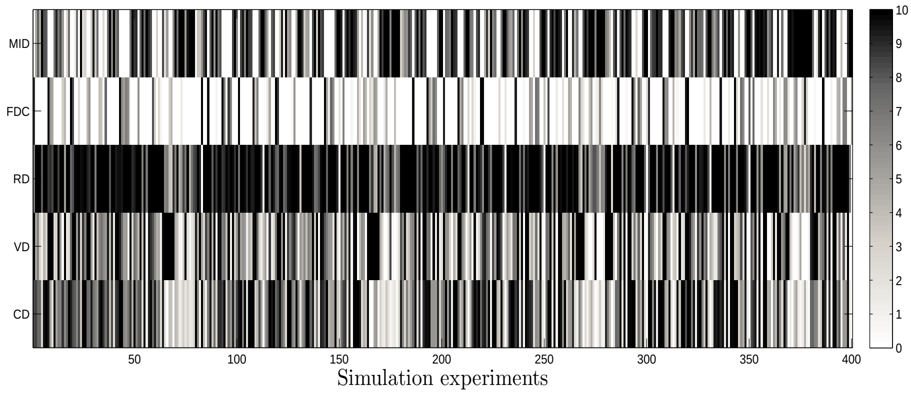

Figure 2b.- Cromatic version of the global relative ranking table taking into account the 400 considered experiments (columns) and the Linear Discriminant Analysis.

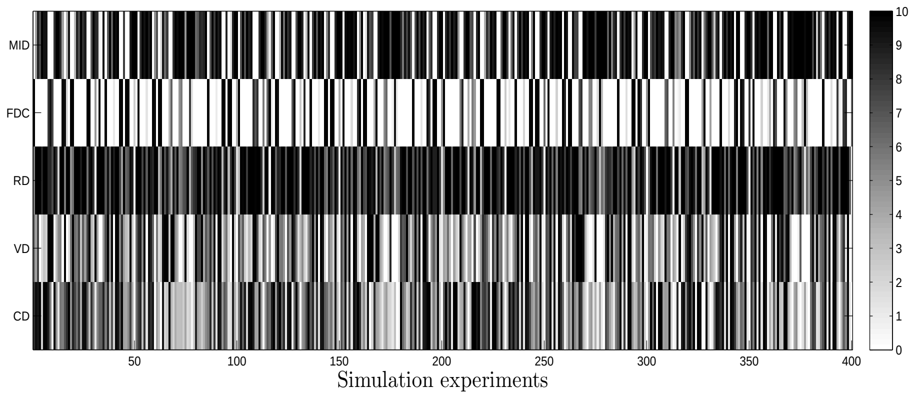

Figure 2c.- Cromatic version of the global relative ranking table taking into account the 400 considered experiments (columns) and the linear SVM.