Supplementary Material:

Band Depth based initialization of k-Means for functional data clustering

Javier Albert-Smet, Aurora Torrente and Juan Romo

Advances in Data Analysis and Classification

1 Results in simulated data: Models 3 and 4

Table S1: Summary statistics for Model 3. The median, mean and variance of 1000 independent datasets for the five performance evaluation measures are provided. Best and second-best medians and means are in boldface and italics, respectively

		Correctness	ARI	Distortion	Iterations	Exec. Time (s)
KM	$\tilde{x}$	0.4400	0.2151	**23680**	*4.000*	**0.0026**
	$\bar{x}$	0.4408	0.2176	**23680**	4.116	**0.0025**
	s	0.0363	0.0377	228.4	0.853	0.0021
KMPP	$\tilde{x}$	0.4400	0.2123	**23680**	*4.000*	*0.0100*
	$\bar{x}$	0.4396	0.2150	**23680**	4.113	*0.0132*
	s	0.0355	0.0360	227.1	0.860	0.0042
BRIk	$\tilde{x}$	0.4320	0.2424	**23680**	*4.000*	0.2445
	$\bar{x}$	0.4361	0.2430	*23690*	4.169	0.2484
	s	0.0324	0.0396	225.3	0.876	0.0221
FKM	$\tilde{x}$	*0.5200*	0.3190	23980	**3.000**	0.2065
	$\bar{x}$	0.5203	0.3183	23980	3.472	0.2110
	s	0.0566	0.0484	230.6	0.693	0.0197
FKMPP	$\tilde{x}$	*0.5200*	0.3106	23980	**3.000**	0.2158
	$\bar{x}$	0.5191	0.3118	23970	3.433	0.2203
	s	0.0562	0.0481	230.4	0.684	0.0203
FABRIk	$\tilde{x}$	**0.5280**	*0.3285*	*23970*	**3.000**	0.4350
	$\bar{x}$	*0.5314*	*0.3283*	23970	**2.603**	0.4415
	s	0.0559	0.0454	229.6	0.556	0.0278
FDEBRIk$_0$	$\tilde{x}$	0.4960	**0.3526**	23980	**3.000**	49.540
	$\bar{x}$	0.5058	**0.3498**	23980	*3.112*	49.750
	s	0.0532	0.0416	229.6	0.720	2.174
FDEBRIk$_1$	$\tilde{x}$	**0.5280**	0.3108	23980	*4.000*	62.720
	$\bar{x}$	**0.5308**	0.3125	23980	3.655	63.030
	s	0.0556	0.0493	230.2	0.746	3.105

Table S2: Summary statistics for Model 3 with 25% missing values. The median, mean and variance of 1000 independent datasets for the five performance evaluation measures are provided. Best and second-best medians and means are in boldface and italics, respectively

		Correctness	ARI	Distortion	Iterations	Exec. Time (s)
KM	$\tilde{x}$	0.4320	0.2010	**20850**	*4.000*	0.6744
	$\bar{x}$	0.4363	0.2030	**20850**	4.091	0.6870
	s	0.0368	0.0363	236.4	0.807	0.0530
KMPP	$\tilde{x}$	0.4320	0.1969	**20850**	*4.000*	0.6835
	$\bar{x}$	0.4356	0.2004	**20850**	4.093	0.6960
	s	0.0358	0.0367	234.6	0.813	0.0532
BRIk	$\tilde{x}$	0.4320	0.2221	*20860*	*4.000*	0.9314
	$\bar{x}$	0.4340	0.2235	*20860*	4.108	0.9473
	s	0.0347	0.0399	237.3	0.883	0.0648
FKM	$\tilde{x}$	*0.4880*	0.2623	21100	**3.000**	**0.2121**
	$\bar{x}$	0.4894	0.2654	21100	3.574	**0.2187**
	s	0.0490	0.0443	240.7	0.720	0.0238
FKMPP	$\tilde{x}$	*0.4880*	0.2629	21100	**3.000**	*0.2202*
	$\bar{x}$	0.4912	0.2680	21100	3.543	*0.2268*
	s	0.0506	0.0443	239.6	0.706	0.0227
FABRIk	$\tilde{x}$	*0.4880*	*0.2785*	21090	**3.000**	0.4511
	$\bar{x}$	*0.4946*	*0.2794*	21090	**2.698**	0.4620
	s	0.0463	0.0421	239.4	0.694	0.0370
FDEBRIk$_0$	$\tilde{x}$	0.4720	**0.3006**	21100	**3.000**	48.900
	$\bar{x}$	0.4809	**0.2989**	21100	*3.184*	49.150
	s	0.0469	0.0408	239.5	0.689	2.957
FDEBRIk$_1$	$\tilde{x}$	**0.4960**	0.2642	21100	*4.000*	62.800
	$\bar{x}$	**0.4976**	0.2662	21100	3.710	63.110
	s	0.0504	0.0440	239.8	0.716	3.124

Table S3: Summary statistics for Model 4. The median, mean and variance of 1000 independent datasets for the five performance evaluation measures are provided. Best and second-best medians and means are in boldface and italics, respectively

		Correctness	ARI	Distortion	Iterations	Exec. Time (s)
KM	$\tilde{x}$	0.5700	0.3070	*1894*	*3.000*	**∼0**
	$\bar{x}$	0.5692	0.3071	*1894*	3.394	**0.0003**
	s	0.0537	0.0558	61.26	0.682	0.0012
KMPP	$\tilde{x}$	0.5800	0.3123	1895	*3.000*	*0.0016*
	$\bar{x}$	0.5719	0.3089	*1894*	3.364	*0.0021*
	s	0.0572	0.0602	60.78	0.716	0.0028
BRIk	$\tilde{x}$	0.5800	0.3176	**1891**	*3.000*	0.0208
	$\bar{x}$	0.5806	0.3183	**1891**	2.667	0.0227
	s	0.0525	0.0568	61.18	0.682	0.0091
FKM	$\tilde{x}$	*0.6200*	0.3678	1938	*3.000*	0.1207
	$\bar{x}$	0.6147	0.3658	1937	2.992	0.1239
	s	0.0538	0.0594	63.27	0.696	0.0180
FKMPP	$\tilde{x}$	*0.6200*	0.3657	1940	*3.000*	0.0932
	$\bar{x}$	0.6110	0.3643	1938	2.902	0.0977
	s	0.0553	0.0616	62.72	0.692	0.0188
FABRIk	$\tilde{x}$	**0.6300**	*0.3730*	1935	**2.000**	0.1128
	$\bar{x}$	**0.6293**	**0.3772**	1933	**2.092**	0.1178
	s	0.0448	0.0560	63.30	0.368	0.0203
FDEBRIk_0	$\tilde{x}$	**0.6300**	0.3727	1940	*3.000*	11.1700
	$\bar{x}$	*0.6287*	*0.3714*	1940	*2.580*	11.3800
	s	0.0488	0.0583	62.98	0.611	1.0210
FDEBRIk_1	$\tilde{x}$	**0.6300**	**0.3733**	1937	*3.000*	17.5700
	$\bar{x}$	0.6243	*0.3714*	1936	3.144	17.5800
	s	0.0494	0.0589	62.78	0.603	0.7596

Table S4: Summary statistics for Model 4 with 25% missing values. The median, mean and variance of 1000 independent datasets for the five performance evaluation measures are provided. Best and second-best medians and means are in boldface and italics, respectively

		Correctness	ARI	Distortion	Iterations	Exec. Time (s)
KM	$\tilde{x}$	0.5700	0.2858	*1653*	*3.000*	**0.0432**
	$\bar{x}$	0.5634	0.2843	*1654*	3.363	**0.0405**
	s	0.0528	0.0567	62.52	0.688	0.0120
KMPP	$\tilde{x}$	0.5600	0.2813	1654	*3.000*	*0.0461*
	$\bar{x}$	0.5626	0.2818	1655	3.329	*0.0421*
	s	0.0518	0.0561	62.00	0.661	0.0129
BRIk	$\tilde{x}$	0.5700	0.2896	**1650**	**2.000**	0.0625
	$\bar{x}$	0.5691	0.2897	**1651**	*2.540*	0.0616
	s	0.0460	0.0539	61.90	0.655	0.0141
FKM	$\tilde{x}$	*0.5900*	0.3119	1686	*3.000*	0.1211
	$\bar{x}$	0.5837	0.3122	1687	3.043	0.1248
	s	0.0503	0.0584	64.06	0.717	0.0185
FKMPP	$\tilde{x}$	0.5800	0.3082	1687	*3.000*	0.0937
	$\bar{x}$	0.5802	0.3085	1688	2.981	0.0938
	s	0.0552	0.0618	63.96	0.675	0.0160
FABRIk	$\tilde{x}$	**0.6000**	0.3238	1682	**2.000**	0.1094
	$\bar{x}$	*0.5959*	0.3205	1683	**2.119**	0.1135
	s	0.0456	0.0567	63.48	0.386	0.0190
FDEBRIk$_0$	$\tilde{x}$	**0.6000**	*0.3246*	1691	*3.000*	12.0600
	$\bar{x}$	**0.5978**	*0.3249*	1691	2.587	12.1000
	s	0.0495	0.0580	63.53	0.647	0.6871
FDEBRIk$_1$	$\tilde{x}$	*0.5900*	**0.3392**	1686	*3.000*	17.7000
	$\bar{x}$	0.5936	**0.3391**	1686	3.234	17.7000
	s	0.0468	0.0561	63.98	0.654	0.7795