

SUPPLEMENTARY MATERIAL FOR JOURNAL ARTICLE

Regularised feed forward neural networks for streamed data classification problems

By Mathys Ellis^{a,}, Anna S. Bosman^b, and Andries P. Engelbrecht^{c,d}*

^aDepartment of Computer Science, University of Pretoria, Pretoria, South Africa

^bDepartment of Computer Science, University of Pretoria, Pretoria, South Africa

^cDepartment of Industrial Engineering and Computer Science Division, Stellenbosch University, Stellenbosch, South Africa

^dCenter for Applied Mathematics and Bioinformatics, Gulf University for Science and Technology, Kuwait

**Corresponding author*

Email address: mox.1990@gmail.com

Appendix A. Empirical evaluation results tables

Tables A.8 to A.10 presents the descriptive statistics and MWU-based ranking results for the various performance metrics of the classifiers, aggregated by classifier-domain, classifier-difficulty, and classifier-environment levels.

Table A.11 presents the MWU-based pairwise comparison of the various classifiers for each of the performance metrics.

Table A.8: Descriptive statistics and MWU-based ranking results of the classifiers aggregated by the problem domains

(a) Training mean square error ($MS E_t$)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.0027±0.0035 1(70.00/27.50/2.50)	0.0027±0.0033 2(70.00/25.00/5.00)	0.0029±0.0036 3(63.75/20.00/16.25)	0.3537±0.0249 6(0.00/0.00/100.00)	0.0664±0.0144 5(20.00/0.00/80.00)	0.0404±0.0173 4(40.00/0.00/60.00)
Sphere	0.0216±0.0153 4(32.50/26.25/41.25)	0.0221±0.0157 5(21.25/18.75/60.00)	0.0216±0.0156 3(36.25/27.50/36.25)	0.3245±0.1101 6(0.00/5.00/95.00)	0.0061±0.0026 2(70.00/5.00/25.00)	0.0142±0.0081 1(97.50/2.50/0.00)
Thresholds	0.0103±0.0056 1(86.25/13.75/0.00)	0.0107±0.0056 2(68.75/20.00/11.25)	0.0113±0.0056 3(55.00/12.50/32.50)	0.2464±0.0451 6(0.00/0.00/100.00)	0.0230±0.0058 5(20.00/0.00/80.00)	0.0143±0.0051 4(40.00/13.75/46.25)
SEA	0.0598±0.0234 2(36.25/32.50/31.25)	0.0606±0.0232 4(25.00/21.25/53.75)	0.0598±0.0235 2(36.25/32.50/31.25)	0.3183±0.0628 5(0.00/0.00/100.00)	0.0043±0.0033 1(100.00/0.00/0.00)	0.0949±0.0400 3(53.75/11.25/35.00)
Electricity	0.0260±0.0182 2(67.50/1.25/31.25)	0.0285±0.0180 5(43.75/1.25/55.00)	0.0536±0.0651 3(60.00/0.00/40.00)	0.4993±0.0920 6(0.00/0.00/100.00)	0.0178±0.0056 1(80.00/0.00/20.00)	0.0279±0.0086 4(46.25/2.50/51.25)

(b) Memory mean square error ($MS E_m$)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.0443±0.0274 2(77.50/11.25/11.25)	0.0455±0.0265 3(60.00/7.50/32.50)	0.0430±0.0254 1(91.25/3.75/5.00)	0.3845±0.0436 6(0.00/0.00/100.00)	0.1448±0.0175 5(20.00/6.25/73.75)	0.1405±0.0186 4(33.75/6.25/60.00)
Sphere	0.0702±0.0335 1(73.75/26.25/0.00)	0.0715±0.0334 3(60.00/26.25/13.75)	0.0710±0.0337 2(65.00/30.00/5.00)	0.3988±0.0476 6(0.00/0.00/100.00)	0.1961±0.0225 5(20.00/1.25/78.75)	0.1813±0.0304 4(38.75/1.25/60.00)
Thresholds	0.0331±0.0174 1(76.25/23.75/0.00)	0.0335±0.0173 2(61.25/30.00/8.75)	0.0337±0.0175 3(61.25/28.75/10.00)	0.2846±0.0163 6(0.00/0.00/100.00)	0.1275±0.0108 5(20.00/5.00/75.00)	0.1098±0.0164 4(35.00/5.00/60.00)
SEA	0.1282±0.0258 1(72.50/27.50/0.00)	0.1291±0.0246 3(61.25/22.50/16.25)	0.1286±0.0257 2(67.50/27.50/5.00)	0.3708±0.0158 6(2.50/5.00/92.50)	0.3603±0.0302 5(12.50/5.00/82.50)	0.2282±0.0286 4(40.00/0.00/60.00)
Electricity	0.2203±0.0436 1(92.50/1.25/6.25)	0.2264±0.0424 2(75.00/1.25/23.75)	0.2901±0.0445 4(37.50/0.00/62.50)	0.5398±0.0409 6(0.00/0.00/100.00)	0.2864±0.0697 5(33.75/0.00/66.25)	0.2714±0.0668 3(60.00/0.00/40.00)

(c) Memory percentage correct classification error (PCC_m)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	80.5178±10.2529 1(85.00/11.25/3.75)	79.5854±9.5886 3(62.50/7.50/30.00)	79.9934±9.4008 2(78.75/8.75/12.50)	61.5537±4.3628 4(40.00/0.00/60.00)	1.5456±0.9681 6(0.00/0.00/100.00)	13.7709±8.1818 5(20.00/0.00/80.00)
Sphere	67.5005±11.7622 1(73.75/16.25/10.00)	66.5257±11.2983 3(46.25/13.75/40.00)	67.0630±11.6111 2(60.00/16.25/23.75)	60.1143±4.7564 5(32.50/11.25/56.25)	55.2463±2.4030 6(15.00/1.25/83.75)	58.4173±2.7255 4(41.25/3.75/55.00)
Thresholds	67.2708±10.9450 1(93.75/6.25/0.00)	66.2988±10.5605 2(75.00/10.00/15.00)	65.1331±10.5987 3(58.75/3.75/37.50)	26.8981±13.7610 5(11.25/11.25/77.50)	24.7291±1.1552 6(7.50/5.00/87.50)	42.4968±6.3503 4(33.75/3.75/62.50)
SEA	35.8517±9.9392 4(20.00/16.25/63.75)	35.2188±9.6418 6(0.00/13.75/86.25)	35.7408±10.0975 5(16.25/17.50/66.25)	62.9222±1.5763 1(86.25/8.75/5.00)	48.3348±3.1967 3(60.00/0.00/40.00)	62.0699±3.5445 2(85.00/8.75/6.25)
Electricity	40.0129±16.3249 1(87.50/1.25/11.25)	35.4457±12.6871 3(67.50/1.25/31.25)	6.6181±7.9148 6(6.25/1.25/92.50)	46.0194±4.0921 2(83.75/0.00/16.25)	14.7658±2.4112 4(37.50/1.25/61.25)	7.4306±1.4391 5(15.00/0.00/85.00)

(d) Generalisation percentage correct classification error (PCC_g)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	81.8334±11.9262 (185.00/12.50/2.50)	81.2043±11.4578 3(61.25/13.75/25.00)	81.5766±11.3502 2(73.75/13.75/12.50)	61.9385±5.0492 4(40.00/0.00/60.00)	1.4871±0.9773 6(0.00/0.00/100.00)	13.5273±8.7044 5(20.00/0.00/80.00)
Sphere	67.5769±12.1462 2(52.50/25.00/22.50)	66.7762±11.9000 4(30.00/20.00/50.00)	67.2979±12.1965 3(42.50/26.25/31.25)	62.4632±6.0127 5(26.25/13.75/60.00)	64.5992±7.9827 6(30.00/2.50/67.50)	69.3535±9.2659 (168.75/12.50/18.75)
Thresholds	66.5919±10.9518 (191.25/7.50/1.25)	65.6423±10.5532 2(75.00/11.25/13.75)	64.4263±10.5758 3(58.75/3.75/37.50)	28.2938±14.3404 5(13.75/8.75/77.50)	25.7685±1.7229 6(7.50/5.00/87.50)	48.9703±10.7458 4(33.75/3.75/62.50)
SEA	35.0198±9.7107 4(17.50/22.50/60.00)	34.4524±9.4374 6(0.00/20.00/80.00)	34.8365±9.6887 5(8.75/25.00/66.25)	64.3005±2.8193 3(67.50/6.25/26.25)	69.6214±16.6769 2(82.50/2.50/15.00)	68.8185±7.0830 (183.75/3.75/12.50)
Electricity	45.1450±17.8577 (188.75/1.25/10.00)	39.0552±15.3828 3(66.25/2.50/31.25)	10.1391±11.4931 5(12.50/1.25/86.25)	47.4473±4.7624 2(82.50/0.00/17.50)	17.8350±2.4587 4(35.00/0.00/65.00)	8.3214±1.5849 6(12.50/0.00/87.50)

(e) Saturation in hidden neurons based on the generalisation set (φ_g)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.6774±0.0707 (188.75/3.75/7.50)	0.6996±0.0860 3(61.25/7.50/31.25)	0.6787±0.0519 2(76.25/6.25/17.50)	0.9978±0.0022 6(0.00/0.00/100.00)	0.7872±0.0095 4(31.25/8.75/60.00)	0.7924±0.0189 5(25.00/8.75/66.25)
Sphere	0.6843±0.0667 (173.75/23.75/2.50)	0.6947±0.0716 3(58.75/21.25/20.00)	0.6864±0.0708 2(72.50/23.75/3.75)	0.9988±0.0016 6(0.00/0.00/100.00)	0.7815±0.0107 4(40.00/1.25/58.75)	0.8524±0.0401 5(20.00/0.00/80.00)
Thresholds	0.6548±0.0898 (172.50/21.25/6.25)	0.6588±0.0930 2(68.75/20.00/11.25)	0.7046±0.1173 4(45.00/8.75/46.25)	0.9967±0.0036 6(0.00/0.00/100.00)	0.7135±0.0063 3(62.50/3.75/33.75)	0.8026±0.0179 5(23.75/1.25/75.00)
SEA	0.7072±0.0801 (155.00/28.75/16.25)	0.7137±0.0799 4(43.75/25.00/31.25)	0.7105±0.0869 2(58.75/21.25/20.00)	0.9987±0.0015 6(0.00/0.00/100.00)	0.7542±0.0037 5(46.25/20.00/33.75)	0.7565±0.0111 3(27.50/42.50/30.00)
Electricity	0.8389±0.0339 4(36.25/3.75/60.00)	0.8615±0.0472 5(20.00/3.75/76.25)	0.4093±0.2423 (197.50/0.00/2.50)	0.9996±0.0005 6(0.00/0.00/100.00)	0.7236±0.0049 2(81.25/0.00/18.75)	0.7376±0.0067 3(61.25/0.00/38.75)

(f) Saturation in hidden neurons based on the memory set (φ_m)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.6777±0.0707 (187.50/5.00/7.50)	0.6999±0.0859 3(61.25/8.75/30.00)	0.6789±0.0518 2(76.25/6.25/17.50)	0.9980±0.0021 6(0.00/0.00/100.00)	0.7872±0.0096 4(31.25/8.75/60.00)	0.7918±0.0189 5(25.00/8.75/66.25)
Sphere	0.6853±0.0671 (173.75/23.75/2.50)	0.6953±0.0716 3(58.75/21.25/20.00)	0.6877±0.0716 2(72.50/23.75/3.75)	0.9989±0.0014 6(0.00/0.00/100.00)	0.7823±0.0096 4(40.00/1.25/58.75)	0.8558±0.0385 5(20.00/0.00/80.00)
Thresholds	0.6679±0.0970 (170.00/23.75/6.25)	0.6700±0.0990 2(66.25/22.50/11.25)	0.7162±0.1236 4(45.00/11.25/43.75)	0.9969±0.0033 6(0.00/0.00/100.00)	0.7147±0.0061 3(62.50/6.25/31.25)	0.8098±0.0193 5(23.75/1.25/75.00)
SEA	0.7079±0.0794 (156.25/27.50/16.25)	0.7144±0.0792 3(45.00/25.00/30.00)	0.7113±0.0864 2(60.00/20.00/20.00)	0.9989±0.0013 6(0.00/0.00/100.00)	0.7582±0.0053 4(53.75/12.50/33.75)	0.7683±0.0107 5(26.25/32.50/41.25)
Electricity	0.8466±0.0329 4(36.25/3.75/60.00)	0.8689±0.0452 5(20.00/3.75/76.25)	0.4093±0.2424 (197.50/0.00/2.50)	0.9996±0.0004 6(0.00/0.00/100.00)	0.7237±0.0048 2(80.00/1.25/18.75)	0.7378±0.0067 3(61.25/1.25/37.50)

(g) Saturation in hidden neurons based on the generalisation set and upper bound of 1 ($\varphi_{0.1,1,g}$)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.6806±0.0676 (187.50/5.00/7.50)	0.7028±0.0830 3(61.25/8.75/30.00)	0.6814±0.0496 2(76.25/6.25/17.50)	0.9996±0.0004 6(0.00/0.00/100.00)	0.7872±0.0095 4(31.25/8.75/60.00)	0.7924±0.0189 5(25.00/8.75/66.25)
Sphere	0.6868±0.0659 (172.50/25.00/2.50)	0.6968±0.0704 3(58.75/23.75/17.50)	0.6894±0.0706 2(71.25/25.00/3.75)	0.9997±0.0003 6(0.00/0.00/100.00)	0.7846±0.0098 4(40.00/1.25/58.75)	0.8639±0.0364 5(20.00/0.00/80.00)
Thresholds	0.6756±0.0989 (167.50/23.75/8.75)	0.6762±0.0994 3(58.75/23.75/17.50)	0.7273±0.1287 3(43.75/13.75/42.50)	0.9989±0.0012 5(0.00/0.00/100.00)	0.7186±0.0061 2(65.00/3.75/31.25)	0.8394±0.0235 4(22.50/2.50/75.00)
SEA	0.7103±0.0780 (165.00/22.50/12.50)	0.7165±0.0780 3(53.75/18.75/27.50)	0.7136±0.0852 2(67.50/18.75/13.75)	0.9997±0.0003 6(0.00/0.00/100.00)	0.7665±0.0088 4(61.25/5.00/33.75)	0.8764±0.0143 5(20.00/0.00/80.00)
Electricity	0.8390±0.0339 4(41.25/3.75/55.00)	0.8616±0.0472 5(22.50/5.00/72.50)	0.4603±0.3042 2(81.25/1.25/17.50)	0.9999±0.0001 6(0.00/0.00/100.00)	0.7237±0.0049 (185.00/0.00/15.00)	0.7376±0.0067 3(65.00/0.00/35.00)

(h) Saturation in hidden neurons based on the generalisation set and upper bound of 5 ($\varphi_{0.1,5,g}$)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.9098±0.0265 (187.50/5.00/7.50)	0.9167±0.0312 3(62.50/10.00/27.50)	0.9103±0.0210 2(78.75/2.50/18.75)	0.9986±0.0014 6(0.00/0.00/100.00)	0.9547±0.0028 5(25.00/7.50/67.50)	0.9536±0.0048 4(30.00/7.50/62.50)
Sphere	0.9012±0.0212 2(72.50/25.00/2.50)	0.9082±0.0243 3(52.50/16.25/31.25)	0.9013±0.0208 (175.00/22.50/2.50)	0.9993±0.0009 6(0.00/0.00/100.00)	0.9247±0.0074 5(32.50/7.50/60.00)	0.9381±0.0284 4(22.50/18.75/58.75)
Thresholds	0.8488±0.0247 (183.75/16.25/0.00)	0.8547±0.0267 2(76.25/16.25/7.50)	0.8715±0.0345 3(51.25/16.25/32.50)	0.9976±0.0025 6(0.00/0.00/100.00)	0.9120±0.0054 5(20.00/10.00/70.00)	0.8948±0.0113 4(31.25/16.25/52.50)
SEA	0.9102±0.0339 2(45.00/23.75/31.25)	0.9138±0.0333 5(25.00/20.00/55.00)	0.9111±0.0340 3(36.25/27.50/36.25)	0.9992±0.0009 6(0.00/0.00/100.00)	0.8907±0.0136 4(57.50/5.00/37.50)	0.7939±0.0114 (196.25/3.75/0.00)
Electricity	0.9647±0.0085 4(38.75/1.25/60.00)	0.9699±0.0108 5(20.00/1.25/78.75)	0.7677±0.0754 (100.00/0.00/0.00)	0.9998±0.0003 6(0.00/0.00/100.00)	0.9235±0.0027 2(80.00/0.00/20.00)	0.9339±0.0034 3(60.00/0.00/40.00)

(i) Saturation in hidden neurons based on the generalisation set and upper bound of 10 ($\varphi_{0.1,10_g}$)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	0.9549±0.0133 1(87.50/5.00/7.50)	0.9583±0.0156 3(62.50/10.00/27.50)	0.9552±0.0105 2(78.75/2.50/18.75)	0.9980±0.0020 6(0.00/0.00/100.00)	0.9773±0.0014 5(25.00/7.50/67.50)	0.9768±0.0024 4(30.00/7.50/62.50)
Sphere	0.9506±0.0106 2(72.50/25.00/2.50)	0.9541±0.0122 3(52.50/16.25/31.25)	0.9507±0.0104 1(75.00/22.50/2.50)	0.9990±0.0013 6(0.00/0.00/100.00)	0.9623±0.0037 5(32.50/7.50/60.00)	0.9632±0.0147 4(22.50/18.75/58.75)
Thresholds	0.9241±0.0121 1(73.75/20.00/6.25)	0.9272±0.0133 2(65.00/21.25/13.75)	0.9348±0.0163 4(48.75/11.25/40.00)	0.9977±0.0025 6(0.00/0.00/100.00)	0.9559±0.0027 5(20.00/5.00/75.00)	0.9265±0.0197 3(55.00/17.50/27.50)
SEA	0.9551±0.0170 2(45.00/22.50/32.50)	0.9569±0.0167 5(22.50/20.00/57.50)	0.9556±0.0170 4(36.25/26.25/37.50)	0.9990±0.0012 6(0.00/0.00/100.00)	0.9449±0.0071 3(58.75/6.25/35.00)	0.8586±0.0194 1(100.00/0.00/0.00)
Electricity	0.9824±0.0043 4(38.75/1.25/60.00)	0.9850±0.0054 5(20.00/1.25/78.75)	0.8838±0.0377 1(100.00/0.00/0.00)	0.9997±0.0004 6(0.00/0.00/100.00)	0.9617±0.0014 2(80.00/0.00/20.00)	0.9669±0.0017 3(60.00/0.00/40.00)

(j) Synapse oversize ratio (n_{sor})

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	1.0353±0.4076 4(32.50/11.25/56.25)	0.9993±0.4344 2(48.75/12.50/38.75)	1.0882±0.3281 6(28.75/10.00/61.25)	0.4831±0.4538 1(92.50/1.25/6.25)	1.0830±0.0758 5(27.50/13.75/58.75)	1.0455±0.1257 3(38.75/13.75/47.50)
Sphere	1.3240±0.1977 6(0.00/17.50/82.50)	1.3182±0.2016 5(3.75/18.75/77.50)	1.2483±0.2260 4(36.25/3.75/60.00)	0.4353±0.3242 1(92.50/5.00/2.50)	0.8249±0.0937 3(61.25/7.50/31.25)	0.7404±0.1358 2(76.25/7.50/16.25)
Thresholds	1.4926±0.1906 5(2.50/23.75/73.75)	1.4923±0.1923 6(1.25/23.75/75.00)	1.3266±0.3675 4(31.25/13.75/55.00)	0.5722±0.2152 1(95.00/5.00/0.00)	0.9077±0.0542 3(56.25/15.00/28.75)	0.8847±0.0706 2(63.75/18.75/17.50)
SEA	1.6581±0.1540 5(2.50/18.75/78.75)	1.6638±0.1553 6(1.25/17.50/81.25)	1.5526±0.1951 4(36.25/3.75/60.00)	0.3129±0.2499 1(97.50/2.50/0.00)	0.8123±0.1015 2(67.50/13.75/18.75)	0.8336±0.1582 3(60.00/13.75/26.25)
Electricity	1.0709±0.3756 6(0.00/6.25/93.75)	0.9983±0.4437 5(21.25/7.50/71.25)	0.2669±0.1122 2(83.75/0.00/16.25)	0.2473±0.2809 1(93.75/2.50/3.75)	0.7608±0.2385 4(35.00/3.75/61.25)	0.7394±0.2357 3(55.00/2.50/42.50)

(k) Hidden neuron oversize ratio (n_{hor})

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	1.8140±0.1184 4(21.25/16.25/62.50)	1.8151±0.1176 5(12.50/15.00/72.50)	1.8201±0.1115 6(6.25/8.75/85.00)	1.2011±0.2073 1(100.00/0.00/0.00)	1.6515±0.0518 3(60.00/7.50/32.50)	1.6022±0.0859 2(72.50/7.50/20.00)
Sphere	1.8054±0.0871 5(6.25/27.50/66.25)	1.8057±0.0861 6(5.00/27.50/67.50)	1.7788±0.1111 4(5.00/32.50/62.50)	1.1468±0.1307 1(97.50/2.50/0.00)	1.3932±0.0423 3(60.00/10.00/30.00)	1.3627±0.0579 2(71.25/10.00/18.75)
Thresholds	1.9997±0.0003 6(0.00/26.25/73.75)	1.9996±0.0004 5(1.25/26.25/72.50)	1.8337±0.2273 4(31.25/16.25/52.50)	1.2406±0.1687 1(100.00/0.00/0.00)	1.6366±0.0277 3(56.25/0.00/43.75)	1.5513±0.0292 2(76.25/1.25/22.50)
SEA	1.9765±0.0282 5(2.50/25.00/72.50)	1.9776±0.0264 6(0.00/23.75/76.25)	1.8996±0.0946 4(26.25/13.75/60.00)	1.0949±0.0925 1(100.00/0.00/0.00)	1.3500±0.0442 2(65.00/15.00/20.00)	1.3749±0.0749 3(60.00/15.00/25.00)
Electricity	1.9237±0.1198 6(3.75/7.50/88.75)	1.8465±0.2632 5(11.25/7.50/81.25)	1.3621±0.1805 2(66.25/7.50/26.25)	1.0792±0.1110 1(100.00/0.00/0.00)	1.4457±0.0934 3(48.75/16.25/35.00)	1.4529±0.0979 4(42.50/16.25/41.25)

(l) Complexity reduction (Ω_r)

Domain	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	38.2707±19.5678 4(32.50/10.00/57.50)	39.9280±20.7881 2(47.50/12.50/40.00)	35.7858±15.8544 6(27.50/10.00/62.50)	68.0819±22.4192 1(92.50/2.50/5.00)	37.2063±3.8543 5(28.75/11.25/60.00)	39.2882±6.4000 3(41.25/13.75/45.00)
Sphere	25.0045±8.1887 6(1.25/16.25/82.50)	25.2092±8.3043 5(3.75/17.50/78.75)	28.0618±9.5303 4(36.25/3.75/60.00)	65.5667±13.3104 1(92.50/7.50/0.00)	48.3571±3.9048 3(60.00/8.75/31.25)	51.7793±5.6133 2(75.00/8.75/16.25)
Thresholds	13.4331±7.0227 5(2.50/23.75/73.75)	13.4445±7.0883 6(1.25/23.75/75.00)	21.2988±15.8277 4(32.50/12.50/55.00)	55.3341±9.6979 1(96.25/3.75/0.00)	38.8067±2.2554 3(56.25/7.50/36.25)	40.5509±2.8741 2(70.00/11.25/18.75)
SEA	10.7911±5.8345 5(3.75/17.50/78.75)	10.5770±5.8643 6(1.25/16.25/82.50)	15.5920±8.1741 4(36.25/3.75/60.00)	70.6273±10.1498 1(97.50/2.50/0.00)	49.3925±4.2119 2(68.75/12.50/18.75)	48.2965±6.6059 3(60.00/12.50/27.50)
Electricity	36.9424±16.0253 6(0.00/6.25/93.75)	40.5884±20.0353 5(18.75/6.25/75.00)	74.4535±6.1764 2(82.50/1.25/16.25)	77.9930±12.2729 1(93.75/3.75/2.50)	53.9616±10.4044 4(37.50/5.00/57.50)	54.7438±10.3306 3(53.75/5.00/41.25)

(m) Swarm diversity (D)

Domain	Classifier		
	QPSO-N	QPSO-WD	QPSO-WE
Hyperplane	$\infty \pm \infty$ 3(0.00/0.00/100.00)	0.5515±0.0448 1(100.00/0.00/0.00)	0.6832±0.0669 2(50.00/0.00/50.00)
Sphere	$\infty \pm \infty$ 3(0.00/0.00/100.00)	1.4283±0.3476 2(56.25/6.25/37.50)	1.1051±0.6068 1(87.50/6.25/6.25)
Thresholds	$\infty \pm \infty$ 3(0.00/0.00/100.00)	1.2556±0.1741 1(78.13/12.50/9.38)	1.4449±0.4049 2(59.38/12.50/28.13)
SEA	$\infty \pm \infty$ 3(0.00/0.00/100.00)	3.3617±1.1768 2(50.00/0.00/50.00)	$\infty \pm \infty$ 1(100.00/0.00/0.00)
Electricity	$\infty \pm \infty$ 3(0.00/0.00/100.00)	1.6456±0.3540 2(50.00/0.00/50.00)	1.4172±0.2540 1(100.00/0.00/0.00)

Table A.9: Descriptive statistics and MWU-based ranking results of the classifiers aggregated by the class of problem difficulty**(a) Training mean square error ($MS E_t$)**

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	0.0134±0.0163 1(68.00/13.00/19.00)	0.0145±0.0168 5(44.00/9.00/47.00)	0.0232±0.0513 2(64.00/16.00/20.00)	0.4090±0.0940 6(0.00/0.00/100.00)	0.0188±0.0183 4(54.00/0.00/46.00)	0.0400±0.0542 3(46.00/10.00/44.00)
Moderate-I	0.0332±0.0308 1(50.00/21.00/29.00)	0.0341±0.0311 5(43.00/15.00/42.00)	0.0336±0.0320 3(51.00/16.00/33.00)	0.3686±0.1016 6(0.00/0.00/100.00)	0.0240±0.0250 4(63.00/0.00/37.00)	0.0398±0.0273 2(67.00/0.00/33.00)
Moderate-II	0.0130±0.0139 1(66.00/19.00/15.00)	0.0136±0.0140 2(54.00/15.00/31.00)	0.0135±0.0142 2(54.00/15.00/31.00)	0.3521±0.0956 5(0.00/0.00/100.00)	0.0216±0.0204 3(52.00/1.00/47.00)	0.0376±0.0373 4(47.00/4.00/49.00)
Hard	0.0367±0.0250 1(50.00/28.00/22.00)	0.0374±0.0254 3(42.00/30.00/28.00)	0.0491±0.0401 5(32.00/27.00/41.00)	0.2640±0.1027 6(0.00/4.00/96.00)	0.0295±0.0304 4(63.00/3.00/34.00)	0.0360±0.0175 2(62.00/10.00/28.00)

(b) Memory mean square error ($MS E_m$)

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	0.0672±0.0639 1(86.00/3.00/4.00)	0.0701±0.0661 3(61.00/4.00/35.00)	0.0849±0.0961 2(64.00/7.00/29.00)	0.4165±0.0930 6(0.00/0.00/100.00)	0.1918±0.0842 5(26.00/1.00/73.00)	0.1590±0.0515 4(48.00/1.00/51.00)
Moderate-I	0.1009±0.0862 1(77.00/22.00/1.00)	0.1028±0.0886 3(64.00/18.00/18.00)	0.1103±0.1039 2(63.00/24.00/13.00)	0.3997±0.0965 6(0.00/1.00/99.00)	0.2283±0.1021 5(20.00/4.00/76.00)	0.1809±0.0713 4(40.00/3.00/57.00)
Moderate-II	0.0957±0.0661 1(77.00/19.00/4.00)	0.0984±0.0689 2(66.00/19.00/15.00)	0.1236±0.1174 3(61.00/18.00/21.00)	0.3873±0.0924 6(0.00/1.00/99.00)	0.2185±0.0908 5(23.00/4.00/73.00)	0.1861±0.0649 4(41.00/3.00/56.00)
Hard	0.1330±0.0744 1(67.00/28.00/5.00)	0.1336±0.0745 3(63.00/29.00/8.00)	0.1343±0.0775 2(70.00/23.00/7.00)	0.3792±0.0788 6(2.00/2.00/96.00)	0.2534±0.1009 5(16.00/5.00/79.00)	0.2189±0.0760 4(37.00/3.00/60.00)

(c) Memory percentage correct classification error (PCC_m)

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	71.1375±17.3775 1(86.00/0.00/14.00)	67.5381±18.8088 2(56.00/1.00/43.00)	61.4998±28.4872 3(52.00/2.00/46.00)	46.4666±19.9619 5(38.00/1.00/61.00)	30.5553±21.3814 6(25.00/1.00/74.00)	41.3074±21.8379 4(40.00/1.00/59.00)
Moderate-I	54.2860±23.0222 1(71.00/12.00/17.00)	53.4626±23.4110 2(51.00/7.00/42.00)	48.5740±30.6027 4(47.00/8.00/45.00)	50.9033±14.5641 3(49.00/9.00/42.00)	28.1327±20.0461 6(21.00/2.00/77.00)	35.9494±24.7014 5(40.00/4.00/56.00)
Moderate-II	63.1757±15.6818 1(81.00/5.00/14.00)	61.6657±15.8855 2(55.00/5.00/40.00)	53.7555±27.3672 3(46.00/8.00/46.00)	51.9559±15.4558 4(45.00/7.00/48.00)	29.8271±21.6721 6(22.00/2.00/76.00)	37.8279±23.6618 5(35.00/5.00/60.00)
Hard	44.3238±18.3547 2(50.00/24.00/26.00)	43.7930±18.4561 3(39.00/24.00/37.00)	39.8093±24.1654 4(31.00/20.00/49.00)	56.6804±9.2216 1(71.00/8.00/21.00)	27.1823±19.9601 6(28.00/1.00/71.00)	32.2637±23.2768 5(41.00/3.00/56.00)

(d) Generalisation mean square error ($MS E_g$)

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	0.0291±0.0260 1(82.00/7.00/11.00)	0.0305±0.0265 3(53.00/4.00/43.00)	0.0389±0.0531 2(73.00/9.00/18.00)	0.4109±0.0936 6(0.00/0.00/100.00)	0.0515±0.0202 4(40.00/4.00/56.00)	0.0628±0.0535 5(36.00/8.00/56.00)
Moderate-I	0.0790±0.0486 1(71.00/22.00/7.00)	0.0798±0.0488 3(57.00/19.00/24.00)	0.0791±0.0496 2(66.00/23.00/11.00)	0.3960±0.0932 6(0.00/2.00/98.00)	0.1593±0.0870 5(26.00/3.00/71.00)	0.1298±0.0458 4(42.00/7.00/51.00)
Moderate-II	0.0358±0.0258 1(79.00/19.00/2.00)	0.0366±0.0261 2(63.00/16.00/21.00)	0.0386±0.0290 3(60.00/19.00/21.00)	0.3606±0.0939 6(0.00/0.00/100.00)	0.0595±0.0206 5(33.00/2.00/65.00)	0.0678±0.0383 4(33.00/8.00/59.00)
Hard	0.1146±0.0630 1(69.00/30.00/1.00)	0.1151±0.0631 2(63.00/32.00/5.00)	0.1196±0.0689 3(58.00/27.00/15.00)	0.3619±0.0844 6(0.00/2.00/98.00)	0.1950±0.0935 5(20.00/4.00/76.00)	0.1728±0.0735 4(40.00/5.00/55.00)

(e) Generalisation percentage correct classification error (PCC_g)

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	70.6192±18.8573 1(82.00/7.00/11.00)	66.6190±21.6140 3(49.00/4.00/47.00)	62.7592±27.7451 2(51.00/7.00/42.00)	47.0649±20.0484 5(34.00/0.00/66.00)	40.2676±32.1852 5(34.00/0.00/66.00)	49.2104±27.8631 4(39.00/4.00/57.00)
Moderate-I	54.2417±21.3290 1(69.00/12.00/19.00)	53.1408±21.9535 2(49.00/10.00/41.00)	47.6658±29.8415 4(43.00/9.00/48.00)	51.2764±14.4195 3(46.00/9.00/45.00)	31.6236±23.4974 6(25.00/2.00/73.00)	37.9179±26.4395 5(45.00/4.00/51.00)
Moderate-II	67.8696±15.6180 1(69.00/15.00/16.00)	66.2484±15.6093 2(51.00/15.00/34.00)	56.9481±26.6778 3(37.00/18.00/45.00)	54.7571±15.8841 5(33.00/7.00/60.00)	40.3784±32.2936 6(33.00/3.00/64.00)	45.1703±29.9416 4(46.00/4.00/50.00)
Hard	44.2032±17.6472 2(47.00/23.00/30.00)	43.6963±17.7294 3(37.00/25.00/38.00)	39.2478±23.7219 5(26.00/22.00/52.00)	58.4563±9.4052 1(71.00/7.00/22.00)	31.1794±23.9303 6(32.00/3.00/65.00)	34.8941±25.8256 4(45.00/4.00/51.00)

(f) Saturation in hidden neurons based on the generalisation set (φ_g)

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	0.7753±0.0546 2(67.00/9.00/24.00)	0.8041±0.0664 4(41.00/12.00/47.00)	0.7538±0.0695 3(59.00/10.00/31.00)	0.9998±0.0001 6(0.00/0.00/100.00)	0.7510±0.0307 1(72.00/5.00/23.00)	0.8048±0.0595 5(36.00/14.00/50.00)
Moderate-I	0.7564±0.0619 2(60.00/21.00/19.00)	0.7644±0.0666 3(47.00/21.00/32.00)	0.6288±0.2396 1(65.00/17.00/18.00)	0.9992±0.0006 6(0.00/0.00/100.00)	0.7563±0.0323 4(55.00/9.00/36.00)	0.7916±0.0445 5(31.00/16.00/53.00)
Moderate-II	0.6813±0.0735 2(73.00/10.00/17.00)	0.6939±0.0768 3(57.00/8.00/35.00)	0.6245±0.1326 1(80.00/7.00/13.00)	0.9983±0.0017 6(0.00/0.00/100.00)	0.7502±0.0318 4(42.00/7.00/51.00)	0.7851±0.0446 5(29.00/6.00/65.00)
Hard	0.6371±0.1113 2(61.00/25.00/14.00)	0.6401±0.1128 3(57.00/21.00/22.00)	0.5444±0.1464 1(76.00/14.00/10.00)	0.9960±0.0031 6(0.00/0.00/100.00)	0.7505±0.0300 4(40.00/6.00/54.00)	0.7716±0.0222 5(30.00/6.00/64.00)

(g) Synapse oversize ratio (n_{sor})

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	1.0001±0.4229 6(12.00/10.00/78.00)	0.9489±0.4850 5(26.00/11.00/63.00)	0.8042±0.4267 3(51.00/3.00/46.00)	0.1467±0.1156 1(100.00/0.00/0.00)	0.7546±0.2130 4(38.00/8.00/54.00)	0.6771±0.2024 2(53.00/8.00/39.00)
Moderate-I	1.4224±0.2995 6(4.00/14.00/82.00)	1.4041±0.3392 5(14.00/15.00/71.00)	1.1824±0.5020 4(48.00/3.00/49.00)	0.3447±0.1930 1(100.00/0.00/0.00)	0.9392±0.1134 3(48.00/13.00/39.00)	0.9225±0.1199 2(58.00/11.00/31.00)
Moderate-II	1.2433±0.2548 6(12.00/12.00/76.00)	1.2321±0.2663 5(17.00/12.00/71.00)	1.0308±0.4299 4(43.00/3.00/54.00)	0.2972±0.1617 1(100.00/0.00/0.00)	0.8281±0.1232 3(49.00/5.00/46.00)	0.7993±0.1365 2(60.00/6.00/34.00)
Hard	1.5991±0.1493 6(2.00/26.00/72.00)	1.5924±0.1520 5(4.00/26.00/70.00)	1.3685±0.5401 4(31.00/16.00/53.00)	0.8520±0.2791 1(77.00/13.00/10.00)	0.9891±0.1074 3(63.00/17.00/20.00)	0.9960±0.1106 2(64.00/20.00/16.00)

(h) Hidden neuron oversize ratio (n_{hor})

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	1.8347±0.1388 6(9.00/12.00/79.00)	1.7778±0.2310 5(11.00/12.00/77.00)	1.5523±0.2595 4(41.00/5.00/54.00)	1.0293±0.0276 1(100.00/0.00/0.00)	1.4437±0.1495 3(53.00/6.00/41.00)	1.3910±0.1216 2(65.00/7.00/28.00)
Moderate-I	1.9454±0.0693 6(3.00/22.00/75.00)	1.9407±0.0715 5(6.00/19.00/75.00)	1.8067±0.2033 4(27.00/19.00/54.00)	1.1235±0.0923 1(100.00/0.00/0.00)	1.5309±0.1282 3(60.00/9.00/31.00)	1.5105±0.0923 2(65.00/9.00/26.00)
Moderate-II	1.8673±0.1318 5(13.00/17.00/70.00)	1.8688±0.1309 6(4.00/18.00/78.00)	1.7255±0.2082 4(26.00/13.00/61.00)	1.0999±0.0767 1(100.00/0.00/0.00)	1.4670±0.1288 3(60.00/7.00/33.00)	1.4425±0.1023 2(66.00/7.00/27.00)
Hard	1.9681±0.0492 6(2.00/31.00/67.00)	1.9683±0.0492 5(3.00/31.00/66.00)	1.8708±0.1921 4(14.00/26.00/60.00)	1.3574±0.1520 1(98.00/2.00/0.00)	1.5399±0.1269 3(59.00/17.00/24.00)	1.5312±0.1104 2(62.00/17.00/21.00)

(i) Complexity reduction (Ω_r)

Difficulty	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Easy	38.3735±20.2503 6(12.00/10.00/78.00)	41.0578±23.6296 5(26.00/10.00/64.00)	48.4252±20.5454 3(52.00/2.00/46.00)	79.2003±8.1788 1(100.00/0.00/0.00)	50.8954±11.0611 4(40.00/3.00/57.00)	54.4435±10.4101 2(55.00/5.00/40.00)
Moderate-I	20.3844±13.4807 6(5.00/12.00/83.00)	21.2301±15.1635 5(11.00/14.00/75.00)	31.1449±23.3110 4(47.00/3.00/50.00)	70.4918±10.7328 1(100.00/0.00/0.00)	42.8294±6.4122 3(51.00/8.00/41.00)	43.6601±6.3657 2(63.00/9.00/28.00)
Moderate-II	28.0433±12.4052 6(13.00/11.00/76.00)	28.5030±12.9329 5(17.00/11.00/72.00)	37.7471±20.6881 4(43.00/3.00/54.00)	72.5754±9.5302 1(100.00/0.00/0.00)	47.7451±6.4639 3(49.00/5.00/46.00)	49.0907±6.5124 2(60.00/6.00/34.00)
Hard	12.7523±7.1299 6(2.00/26.00/72.00)	13.0067±7.2815 5(4.00/26.00/70.00)	22.8363±24.8065 4(30.00/17.00/53.00)	47.8150±13.3698 1(78.00/16.00/6.00)	40.7094±5.4741 3(61.00/20.00/19.00)	40.5326±5.3570 2(62.00/21.00/17.00)

(j) Swarm diversity (\mathcal{D})

Difficulty	Classifier		
	QPSO-N	QPSO-WD	QPSO-WE
Easy	∞±∞ 3(0.00/0.00/100.00)	1.2432±0.6265 2(60.00/5.00/35.00)	0.7483±0.3104 1(85.00/5.00/10.00)
Moderate-I	∞±∞ 3(0.00/0.00/100.00)	1.9026±1.3402 2(62.50/5.00/32.50)	∞±∞ 1(82.50/5.00/12.50)
Moderate-II	∞±∞ 3(0.00/0.00/100.00)	1.3394±0.6276 2(70.00/0.00/30.00)	0.9760±0.3721 1(80.00/0.00/20.00)
Hard	∞±∞ 3(0.00/0.00/100.00)	2.1090±1.3496 1(75.00/5.00/20.00)	2.0720±1.3742 2(70.00/5.00/25.00)

Table A.10: Descriptive statistics and MWU-based ranking results of the classifiers aggregated by the class of problem environment**(a) Training mean square error ($MS E_t$)**

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	0.0226±0.0230 1(57.22/26.67/16.11)	0.0230±0.0233 2(50.00/25.56/24.44)	0.0231±0.0229 3(43.33/23.33/33.33)	0.2722±0.0748 6(0.00/2.22/97.78)	0.0264±0.0272 5(50.00/2.22/47.78)	0.0353±0.0311 4(56.11/6.67/37.22)
Abrupt	0.0122±0.0124 1(74.29/8.57/17.14)	0.0136±0.0128 4(45.71/2.86/51.43)	0.0430±0.0832 2(65.71/11.43/22.86)	0.4392±0.0990 6(0.00/0.00/100.00)	0.0160±0.0161 3(60.00/0.00/40.00)	0.0473±0.0641 5(40.00/5.71/54.29)
Chaotic	0.0278±0.0275 1(56.76/16.22/27.03)	0.0289±0.0276 5(41.62/11.89/46.49)	0.0339±0.0389 2(54.05/15.14/30.81)	0.4054±0.0945 6(0.00/0.00/100.00)	0.0221±0.0216 3(65.41/0.00/34.59)	0.0396±0.0346 4(57.84/5.41/36.76)

(b) Memory mean square error ($MS E_m$)

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	0.0789±0.0462 2(67.78/27.78/4.44)	0.0795±0.0461 3(61.11/28.89/10.00)	0.0788±0.0463 1(69.44/26.67/3.89)	0.3431±0.0571 6(1.11/1.67/97.22)	0.2041±0.0940 5(17.22/5.56/77.22)	0.1680±0.0548 4(36.11/3.89/60.00)
Abrupt	0.0721±0.0594 1(91.43/0.00/8.57)	0.0757±0.0612 3(60.00/0.00/40.00)	0.1069±0.1149 2(62.86/0.00/37.14)	0.4430±0.0981 6(0.00/0.00/100.00)	0.1911±0.0875 5(31.43/0.00/68.57)	0.1567±0.0536 4(54.29/0.00/45.71)
Chaotic	0.1241±0.0931 1(86.49/11.89/1.62)	0.1272±0.0950 2(66.49/9.73/23.78)	0.1481±0.1227 3(60.00/12.97/27.03)	0.4379±0.0889 6(0.00/0.54/99.46)	0.2474±0.0949 5(23.24/2.16/74.59)	0.2095±0.0769 4(44.32/1.62/54.05)

(c) Memory percentage correct classification error (PCC_m)

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	59.6057±17.4275 1(63.89/16.67/19.44)	58.8505±17.2130 3(46.67/16.67/36.67)	58.6347±17.0091 2(48.33/15.56/36.11)	52.0496±17.1994 4(45.56/9.44/45.00)	31.3512±20.3961 6(20.00/2.22/77.78)	42.2506±19.5774 5(42.78/5.00/52.22)
Abrupt	69.7218±17.8179 1(77.14/0.00/22.86)	63.9425±19.8207 3(48.57/2.86/48.57)	57.3227±28.4463 2(51.43/5.71/42.86)	46.6844±20.0509 4(45.71/2.86/51.43)	32.7962±22.0601 6(28.57/2.86/68.57)	41.5980±24.1980 5(40.00/2.86/57.14)
Chaotic	54.7189±24.0935 1(78.92/5.95/15.14)	53.0533±24.1710 3(54.05/3.24/42.70)	42.1802±34.6732 4(38.38/4.32/57.30)	51.8796±12.7650 2(56.76/3.78/39.46)	25.8306±20.2804 6(27.03/0.54/72.43)	30.6692±25.2160 5(35.14/1.62/63.24)

(d) Generalisation mean square error ($MS E_g$)

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	0.0592±0.0500 1(70.00/28.33/1.67)	0.0596±0.0499 2(61.11/27.78/11.11)	0.0594±0.0497 3(61.67/26.67/11.67)	0.3231±0.0601 6(0.00/1.11/98.89)	0.1149±0.0887 5(23.33/3.33/73.33)	0.1057±0.0702 4(36.67/7.22/56.11)
Abrupt	0.0252±0.0192 1(85.71/5.71/8.57)	0.0269±0.0196 3(51.43/2.86/45.71)	0.0555±0.0825 2(77.14/5.71/17.14)	0.4398±0.0988 6(0.00/0.00/100.00)	0.0403±0.0161 4(45.71/2.86/51.43)	0.0629±0.0633 5(31.43/0.00/68.57)
Chaotic	0.0773±0.0610 1(78.38/13.51/8.11)	0.0785±0.0610 3(58.38/10.81/30.81)	0.0809±0.0665 2(64.32/15.14/20.54)	0.4291±0.0845 6(0.00/1.08/98.92)	0.1321±0.0925 5(32.97/3.24/63.78)	0.1195±0.0698 4(40.00/8.11/51.89)

(e) Generalisation percentage correct classification error (PCC_g)

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	60.1007±18.5690 1(55.00/21.67/23.33)	59.4296±18.4118 2(42.22/22.78/35.00)	59.1302±18.2265 3(39.44/22.22/38.33)	54.1608±17.7076 5(40.56/8.89/50.56)	38.7415±28.7047 6(29.44/3.89/66.67)	48.2806±24.3954 4(51.11/5.00/43.89)
Abrupt	68.4628±19.4940 1(80.00/2.86/17.14)	61.9730±23.3047 3(40.00/8.57/51.43)	58.9761±26.9168 2(48.57/8.57/42.86)	47.0126±20.1263 4(42.86/0.00/57.14)	44.2696±35.1474 5(40.00/0.00/60.00)	49.2819±30.1378 6(37.14/2.86/60.00)
Chaotic	56.6435±23.4713 1(76.22/8.11/15.68)	54.6165±23.7548 2(51.89/5.41/42.70)	42.9973±33.7676 4(37.30/7.03/55.68)	52.7626±12.6670 3(51.89/3.78/44.32)	31.4702±26.1957 6(30.81/0.54/68.65)	34.0751±28.8328 5(37.84/3.24/58.92)

(f) Saturation in hidden neurons based on the generalisation set (φ_g)

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	0.6303±0.0655 1(75.56/19.44/5.00)	0.6369±0.0682 2(66.67/16.11/17.22)	0.6481±0.0773 3(68.33/12.22/19.44)	0.9968±0.0027 6(0.00/0.00/100.00)	0.7524±0.0317 4(36.11/7.22/56.67)	0.7908±0.0319 5(22.22/7.22/70.56)
Abrupt	0.7932±0.0566 2(65.71/5.71/28.57)	0.8320±0.0724 5(31.43/11.43/57.14)	0.7753±0.0605 3(57.14/5.71/37.14)	0.9999±0.0000 6(0.00/0.00/100.00)	0.7475±0.0289 1(77.14/8.57/14.29)	0.7958±0.0665 4(42.86/20.00/37.14)
Chaotic	0.7773±0.0539 2(55.14/15.14/29.73)	0.7918±0.0584 4(38.38/15.68/45.95)	0.6020±0.2332 1(74.05/12.97/12.97)	0.9995±0.0004 6(0.00/0.00/100.00)	0.7525±0.0309 3(63.24/5.95/30.81)	0.7844±0.0528 5(38.38/11.89/49.73)

(g) Synapse oversize ratio (n_{sor})

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	1.4498±0.2815 6(7.78/21.11/71.11)	1.4457±0.2863 5(9.44/21.11/69.44)	1.3981±0.3051 4(26.67/11.11/62.22)	0.6064±0.3403 1(90.00/6.11/3.89)	0.9098±0.1353 3(59.44/12.78/27.78)	0.9010±0.1566 2(62.78/15.56/21.67)
Abrupt	0.8959±0.4648 6(8.57/8.57/82.86)	0.8327±0.5350 5(25.71/11.43/62.86)	0.6786±0.4619 3(57.14/0.00/42.86)	0.1008±0.0971 1(100.00/0.00/0.00)	0.6469±0.2452 4(37.14/8.57/54.29)	0.5615±0.2032 2(54.29/5.71/40.00)
Chaotic	1.2658±0.3593 6(7.03/11.35/81.62)	1.2345±0.4045 5(18.92/11.89/69.19)	0.8821±0.5267 3(56.76/2.70/40.54)	0.2778±0.2210 1(97.30/1.08/1.62)	0.8902±0.1557 4(42.16/9.19/48.65)	0.8522±0.1697 2(55.68/8.11/36.22)

(h) Hidden neuron oversize ratio (n_{hor})

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	1.9138±0.1156 5(7.22/26.67/66.11)	1.9147±0.1147 6(3.33/27.22/69.44)	1.8729±0.1381 4(12.78/21.11/66.11)	1.2444±0.1748 1(98.89/1.11/0.00)	1.5200±0.1437 3(61.11/8.89/30.00)	1.4892±0.1199 2(69.44/9.44/21.11)
Abrupt	1.8063±0.1611 6(5.71/11.43/82.86)	1.7128±0.3011 5(17.14/11.43/71.43)	1.4687±0.3064 4(45.71/5.71/48.57)	1.0133±0.0134 1(100.00/0.00/0.00)	1.3793±0.1532 3(48.57/8.57/42.86)	1.3276±0.1179 2(60.00/8.57/31.43)
Chaotic	1.9127±0.1023 5(6.49/16.22/77.30)	1.8971±0.1364 6(6.49/14.59/78.92)	1.6595±0.2419 4(37.30/12.43/50.27)	1.0895±0.0891 1(100.00/0.00/0.00)	1.4935±0.1193 3(56.76/10.81/32.43)	1.4756±0.1024 2(60.54/10.81/28.65)

(i) Complexity reduction (Ω_r)

Environment	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
Progressive	18.3308±12.8432 6(8.33/20.56/71.11)	18.4948±13.0873 5(9.44/20.56/70.00)	20.6277±13.5472 4(26.67/11.11/62.22)	57.4849±15.6555 1(90.56/7.22/2.22)	42.8254±6.6542 3(58.33/12.78/28.89)	43.4609±7.0773 2(63.33/14.44/22.22)
Abrupt	42.9507±22.4625 6(8.57/8.57/82.86)	46.4366±26.5322 5(25.71/8.57/65.71)	54.4125±22.7194 3(57.14/0.00/42.86)	81.3608±7.3639 1(100.00/0.00/0.00)	56.2029±12.7666 4(40.00/5.71/54.29)	60.0649±10.8693 2(54.29/5.71/40.00)
Chaotic	27.8515±16.1282 6(7.57/10.27/82.16)	29.3267±18.3145 5(17.30/11.35/71.35)	45.3942±24.4885 3(56.22/2.70/41.08)	74.6667±10.5461 1(97.30/1.62/1.08)	46.1744±7.7276 4(44.32/5.95/49.73)	47.8241±7.9971 2(57.84/7.03/35.14)

(j) Swarm diversity (\mathcal{D})

Environment	Classifier		
	QPSO-N	QPSO-WD	QPSO-WE
Progressive	∞±∞ 3(0.00/0.00/100.00)	1.6429±1.1741 1(76.39/5.56/18.06)	1.4955±1.2028 2(68.06/5.56/26.39)
Abrupt	∞±∞ 3(0.00/0.00/100.00)	1.1877±0.4920 2(57.14/0.00/42.86)	0.6416±0.3054 1(92.86/0.00/7.14)
Chaotic	∞±∞ 3(0.00/0.00/100.00)	1.7412±1.0917 2(59.46/2.70/37.84)	∞±∞ 1(87.84/2.70/9.46)

Table A.11: MWU-based pairwise comparison of the classifiers (Wins/Ties/Losses percentages)**(a) Training mean square error ($MS E_t$)**

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	53.75/45.00/1.25	36.25/48.75/15.00	98.75/1.25/0.00	48.75/0.00/51.25	55.00/6.25/38.75
BP-WD	1.25/45.00/53.75	—	31.25/30.00/38.75	98.75/1.25/0.00	47.50/0.00/52.50	50.00/10.00/40.00
BP-WE	15.00/48.75/36.25	38.75/30.00/31.25	—	98.75/1.25/0.00	51.25/1.25/47.50	47.50/11.25/41.25
QPSO-N	0.00/1.25/98.75	0.00/1.25/98.75	0.00/1.25/98.75	—	0.00/1.25/98.75	0.00/0.00/100.00
QPSO-WD	51.25/0.00/48.75	52.50/0.00/47.50	47.50/1.25/51.25	98.75/1.25/0.00	—	40.00/2.50/57.50
QPSO-WE	38.75/6.25/55.00	40.00/10.00/50.00	41.25/11.25/47.50	100.00/0.00/0.00	57.50/2.50/40.00	—
Overall rank	1(58.50/20.25/21.25)	3(45.75/17.25/37.00)	2(50.25/18.50/31.25)	6(0.00/1.00/99.00)	5(58.00/1.00/41.00)	4(55.50/6.00/38.50)

(b) Memory mean square error ($MS E_m$)

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	55.00/43.75/1.25	41.25/46.25/12.50	100.00/0.00/0.00	98.75/0.00/1.25	97.50/0.00/2.50
BP-WD	1.25/43.75/55.00	—	21.25/43.75/35.00	100.00/0.00/0.00	98.75/0.00/1.25	96.25/0.00/3.75
BP-WE	12.50/46.25/41.25	35.00/43.75/21.25	—	100.00/0.00/0.00	88.75/0.00/11.25	86.25/0.00/13.75
QPSO-N	0.00/0.00/100.00	0.00/0.00/100.00	0.00/0.00/100.00	—	2.50/5.00/92.50	0.00/0.00/100.00
QPSO-WD	1.25/0.00/98.75	1.25/0.00/98.75	11.25/0.00/88.75	92.50/5.00/2.50	—	0.00/12.50/87.50
QPSO-WE	2.50/0.00/97.50	3.75/0.00/96.25	13.75/0.00/86.25	100.00/0.00/0.00	87.50/12.50/0.00	—
Overall rank	1(78.50/18.00/3.50)	3(63.50/17.50/19.00)	2(64.50/18.00/17.50)	6(0.50/1.00/98.50)	5(21.25/3.50/75.25)	4(41.50/2.50/56.00)

(c) Memory percentage correct classification error (PCC_m)

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	77.50/22.50/0.00	68.75/23.75/7.50	61.25/3.75/35.00	76.25/1.25/22.50	76.25/0.00/23.75
BP-WD	0.00/22.50/77.50	—	38.75/20.00/41.25	60.00/3.75/36.25	76.25/0.00/23.75	76.25/0.00/23.75
BP-WE	7.50/23.75/68.75	41.25/20.00/38.75	—	52.50/2.50/45.00	57.50/1.25/41.25	61.25/0.00/38.75
QPSO-N	35.00/3.75/61.25	36.25/3.75/60.00	45.00/2.50/52.50	—	82.50/5.00/12.50	55.00/16.25/28.75
QPSO-WD	22.50/1.25/76.25	23.75/0.00/76.25	41.25/1.25/57.50	12.50/5.00/82.50	—	20.00/0.00/80.00
QPSO-WE	23.75/0.00/76.25	23.75/0.00/76.25	38.75/0.00/61.25	28.75/16.25/55.00	80.00/0.00/20.00	—
Overall rank	1(72.00/10.25/17.75)	2(50.25/9.25/40.50)	4(44.00/9.50/46.50)	3(50.75/6.25/43.00)	6(24.00/1.50/74.50)	5(39.00/3.25/57.75)

(d) Generalisation mean square error ($MS E_g$)

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	60.00/40.00/0.00	40.00/48.75/11.25	100.00/0.00/0.00	88.75/0.00/11.25	87.50/8.75/3.75
BP-WD	0.00/40.00/60.00	—	23.75/38.75/37.50	100.00/0.00/0.00	85.00/1.25/13.75	86.25/8.75/5.00
BP-WE	11.25/48.75/40.00	37.50/38.75/23.75	—	100.00/0.00/0.00	87.50/1.25/11.25	85.00/8.75/6.25
QPSO-N	0.00/0.00/100.00	0.00/0.00/100.00	0.00/0.00/100.00	—	0.00/5.00/95.00	0.00/0.00/100.00
QPSO-WD	11.25/0.00/88.75	13.75/1.25/85.00	11.25/1.25/87.50	95.00/5.00/0.00	—	17.50/8.75/73.75
QPSO-WE	3.75/8.75/87.50	5.00/8.75/86.25	6.25/8.75/85.00	100.00/0.00/0.00	73.75/8.75/17.50	—
Overall rank	1(75.25/19.50/5.25)	3(59.00/17.75/23.25)	2(64.25/19.50/16.25)	6(0.00/1.00/99.00)	5(29.75/3.25/67.00)	4(37.75/7.00/55.25)

(e) Generalisation percentage correct classification error (PCC_g)

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	70.00/30.00/0.00	65.00/32.50/2.50	61.25/2.50/36.25	73.75/0.00/26.25	65.00/3.75/31.25
BP-WD	0.00/30.00/70.00	—	36.25/32.50/31.25	58.75/2.50/38.75	73.75/0.00/26.25	63.75/2.50/33.75
BP-WE	2.50/32.50/65.00	31.25/32.50/36.25	—	51.25/2.50/46.25	58.75/0.00/41.25	52.50/2.50/45.00
QPSO-N	36.25/2.50/61.25	38.75/2.50/58.75	46.25/2.50/51.25	—	58.75/10.00/31.25	50.00/11.25/38.75
QPSO-WD	26.25/0.00/73.75	26.25/0.00/73.75	41.25/0.00/58.75	31.25/10.00/58.75	—	30.00/0.00/70.00
QPSO-WE	31.25/3.75/65.00	33.75/2.50/63.75	45.00/2.50/52.50	38.75/11.25/50.00	70.00/0.00/30.00	—
Overall rank	1(67.00/13.75/19.25)	2(46.50/13.50/40.00)	3(39.25/14.00/46.75)	4(46.00/5.75/48.25)	6(31.00/2.00/67.00)	5(43.75/4.00/52.25)

(f) Saturation in hidden neurons based on the generalisation set (φ_g)

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	56.25/43.75/0.00	36.25/27.50/36.25	100.00/0.00/0.00	63.75/1.25/35.00	70.00/8.75/21.25
BP-WD	0.00/43.75/56.25	—	25.00/23.75/51.25	100.00/0.00/0.00	58.75/2.50/38.75	68.75/7.50/23.75
BP-WE	36.25/27.50/36.25	51.25/23.75/25.00	—	100.00/0.00/0.00	78.75/1.25/20.00	83.75/7.50/8.75
QPSO-N	0.00/0.00/100.00	0.00/0.00/100.00	0.00/0.00/100.00	—	0.00/0.00/100.00	0.00/0.00/100.00
QPSO-WD	35.00/1.25/63.75	38.75/2.50/58.75	20.00/1.25/78.75	100.00/0.00/0.00	—	67.50/28.75/3.75
QPSO-WE	21.25/8.75/70.00	23.75/7.50/68.75	8.75/7.50/83.75	100.00/0.00/0.00	3.75/28.75/67.50	—
Overall rank	2(65.25/16.25/18.50)	3(50.50/15.50/34.00)	1(70.00/12.00/18.00)	6(0.00/0.00/100.00)	4(52.25/6.75/41.00)	5(31.50/10.50/58.00)

(g) Synapse oversize ratio (n_{sor})

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	3.75/63.75/32.50	13.75/11.25/75.00	0.00/0.00/100.00	10.00/1.25/88.75	10.00/1.25/88.75
BP-WD	32.50/63.75/3.75	—	13.75/13.75/72.50	0.00/0.00/100.00	15.00/2.50/82.50	15.00/0.00/85.00
BP-WE	75.00/11.25/13.75	72.50/13.75/13.75	—	3.75/0.00/96.25	33.75/2.50/63.75	31.25/3.75/65.00
QPSO-N	100.00/0.00/0.00	100.00/0.00/0.00	96.25/0.00/3.75	—	88.75/6.25/5.00	86.25/10.00/3.75
QPSO-WD	88.75/1.25/10.00	82.50/2.50/15.00	63.75/2.50/33.75	5.00/6.25/88.75	—	7.50/41.25/51.25
QPSO-WE	88.75/1.25/10.00	85.00/0.00/15.00	65.00/3.75/31.25	3.75/10.00/86.25	51.25/41.25/7.50	—
Overall rank	6(7.50/15.50/77.00)	5(15.25/16.00/68.75)	4(43.25/6.25/50.50)	1(94.25/3.25/2.50)	3(49.50/10.75/39.75)	2(58.75/11.25/30.00)

(h) Hidden neuron oversize ratio (n_{hor})

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	20.00/66.25/13.75	13.75/36.25/50.00	0.00/0.00/100.00	0.00/0.00/100.00	0.00/0.00/100.00
BP-WD	13.75/66.25/20.00	—	13.75/33.75/52.50	0.00/0.00/100.00	1.25/0.00/98.75	1.25/0.00/98.75
BP-WE	50.00/36.25/13.75	52.50/33.75/13.75	—	0.00/0.00/100.00	16.25/3.75/80.00	16.25/5.00/78.75
QPSO-N	100.00/0.00/0.00	100.00/0.00/0.00	100.00/0.00/0.00	—	98.75/1.25/0.00	98.75/1.25/0.00
QPSO-WD	100.00/0.00/0.00	98.75/0.00/1.25	80.00/3.75/16.25	0.00/1.25/98.75	—	11.25/43.75/45.00
QPSO-WE	100.00/0.00/0.00	98.75/0.00/1.25	78.75/5.00/16.25	0.00/1.25/98.75	45.00/43.75/11.25	—
Overall rank	5(6.75/20.50/72.75)	6(6.00/20.00/74.00)	4(27.00/15.75/57.25)	1(99.50/0.50/0.00)	3(58.00/9.75/32.25)	2(64.50/10.00/25.50)

(i) Complexity reduction (Ω_r)

Classifier	Classifier					
	BP-N	BP-WD	BP-WE	QPSO-N	QPSO-WD	QPSO-WE
BP-N	—	6.25/61.25/32.50	13.75/11.25/75.00	0.00/0.00/100.00	10.00/1.25/88.75	10.00/0.00/90.00
BP-WD	32.50/61.25/6.25	—	13.75/13.75/72.50	0.00/0.00/100.00	13.75/0.00/86.25	12.50/1.25/86.25
BP-WE	75.00/11.25/13.75	72.50/13.75/13.75	—	2.50/1.25/96.25	35.00/0.00/65.00	30.00/5.00/65.00
QPSO-N	100.00/0.00/0.00	100.00/0.00/0.00	96.25/1.25/2.50	—	88.75/8.75/2.50	87.50/10.00/2.50
QPSO-WD	88.75/1.25/10.00	86.25/0.00/13.75	65.00/0.00/35.00	2.50/8.75/88.75	—	8.75/35.00/56.25
QPSO-WE	90.00/0.00/10.00	86.25/1.25/12.50	65.00/5.00/30.00	2.50/10.00/87.50	56.25/35.00/8.75	—
Overall rank	6(8.00/14.75/77.25)	5(14.50/15.25/70.25)	4(43.00/6.25/50.75)	1(94.50/4.00/1.50)	3(50.25/9.00/40.75)	2(60.00/10.25/29.75)

(j) Swarm diversity (\mathcal{D})

Classifier	Classifier		
	QPSO-N	QPSO-WD	QPSO-WE
QPSO-N	—	0.00/0.00/100.00	0.00/0.00/100.00
QPSO-WD	100.00/0.00/0.00	—	33.75/7.50/58.75
QPSO-WE	100.00/0.00/0.00	58.75/7.50/33.75	—
Overall rank	3(0.00/0.00/100.00)	2(66.88/3.75/29.38)	1(79.38/3.75/16.88)

Appendix B. Selected performance trend graphs

Figures B.7 to B.11 illustrate the accuracy, complexity, and saturation performance trends of the classifiers on the A1 benchmark problems. Because PCC_g , PCC_m , and Ω_r were in the range $[0, 100]$, the values of φ_g , MSE_g , MSE_m , and MSE_t were linearly scaled to the same range by multiplying their values by 100. Furthermore, φ_m was not used because it was shown in the saturation analysis that φ_m was practically identical to φ_g .

Figures B.17 and B.18 present a comparison between the raw MSE_g and MSE_t trends, and the 3% moving average MSE_g and MSE_t trends for the BP classifiers and QPSO classifiers, respectively. The left-hand side graphs represent the raw trends. The right-hand side graphs represent the moving average trends. The moving average trend graphs are also overlaid with the detection results of O_{MSE_g} . Note that figures B.17 and B.18 include the plus and minus moving average standard deviation bands for the MSE_g trend to provide insight into the stability of the MSE_g trend.

Figures B.17 and B.18 represent the typical behaviour found for the classifiers for the benchmark problems.

Figures B.13 to B.16 present the swarm diversity trend graphs. Note that Figures B.13 to B.16 include the plus and minus moving average standard deviation bands for the swarm diversity trend to provide insight into the stability of the swarm diversity trend.

Figure B.13 presents the swarm diversity trend of the QPSO-N on the hyperplane A1 problems. The swarm diversity trends for QPSO-N took on a similar shape for all problems. Problems with higher temporal severity took longer to approach the extreme levels.

Figures B.14 and B.15 present the swarm diversity trends of the regularised QPSO classifiers for the A1 problems.

Figure B.6 provides the legend for the performance trend graphs in Figures B.7 to B.12.

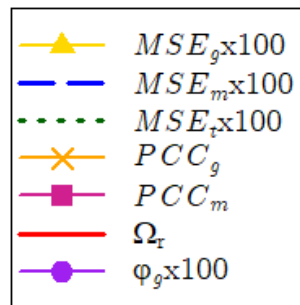


Fig. B.6: Legend for the performance trend analysis graphs

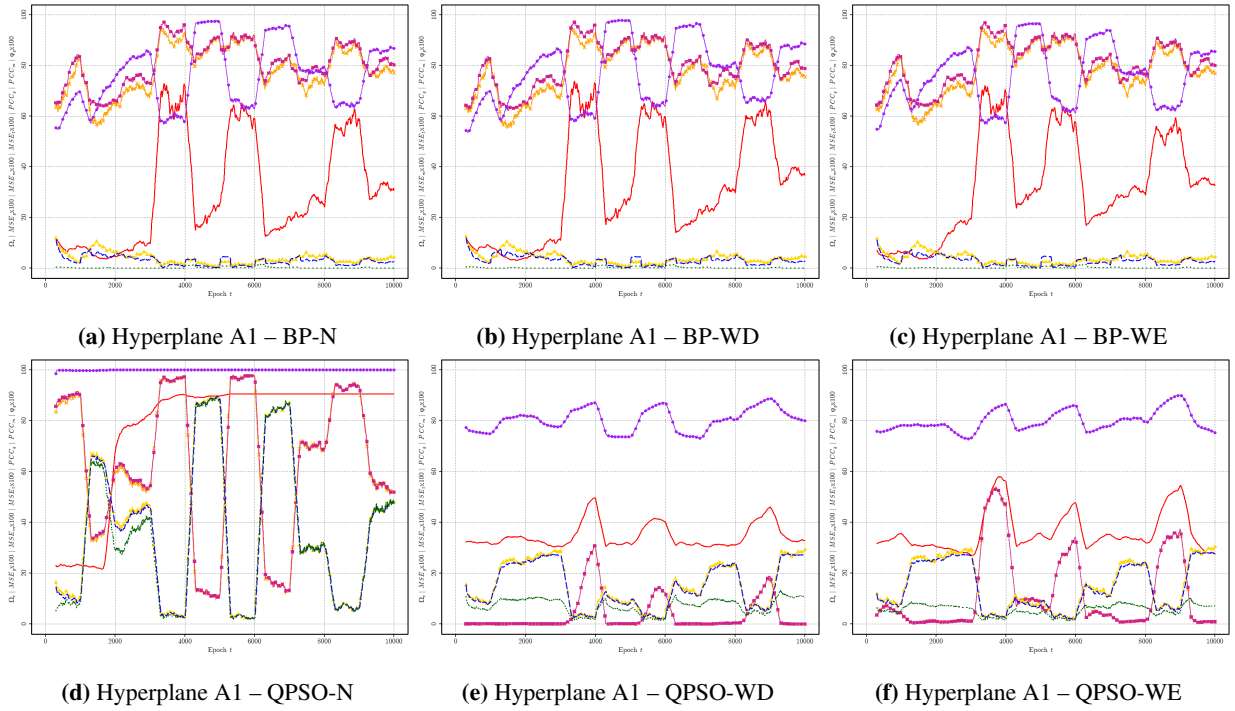


Fig. B.7: Saturation, accuracy, and complexity performance trends of the classifiers for the A1 Hyperplane problems

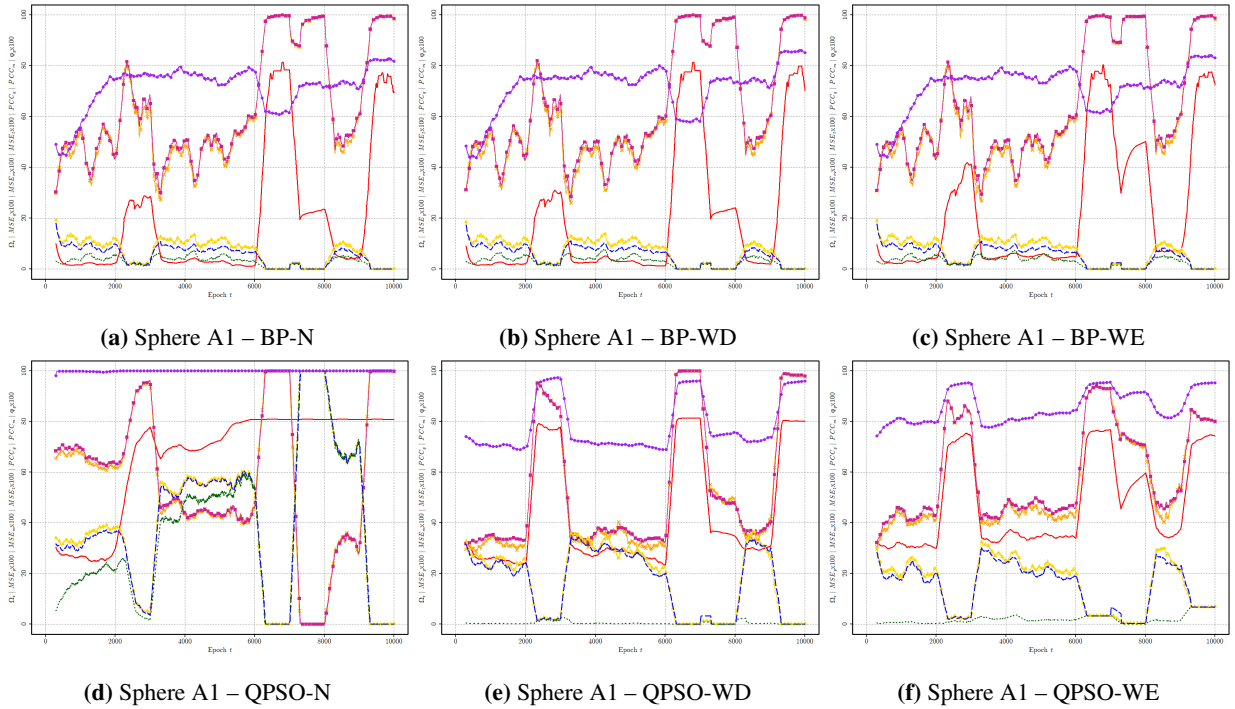


Fig. B.8: Saturation, accuracy, and complexity performance trends of the classifiers for the A1 Sphere problems

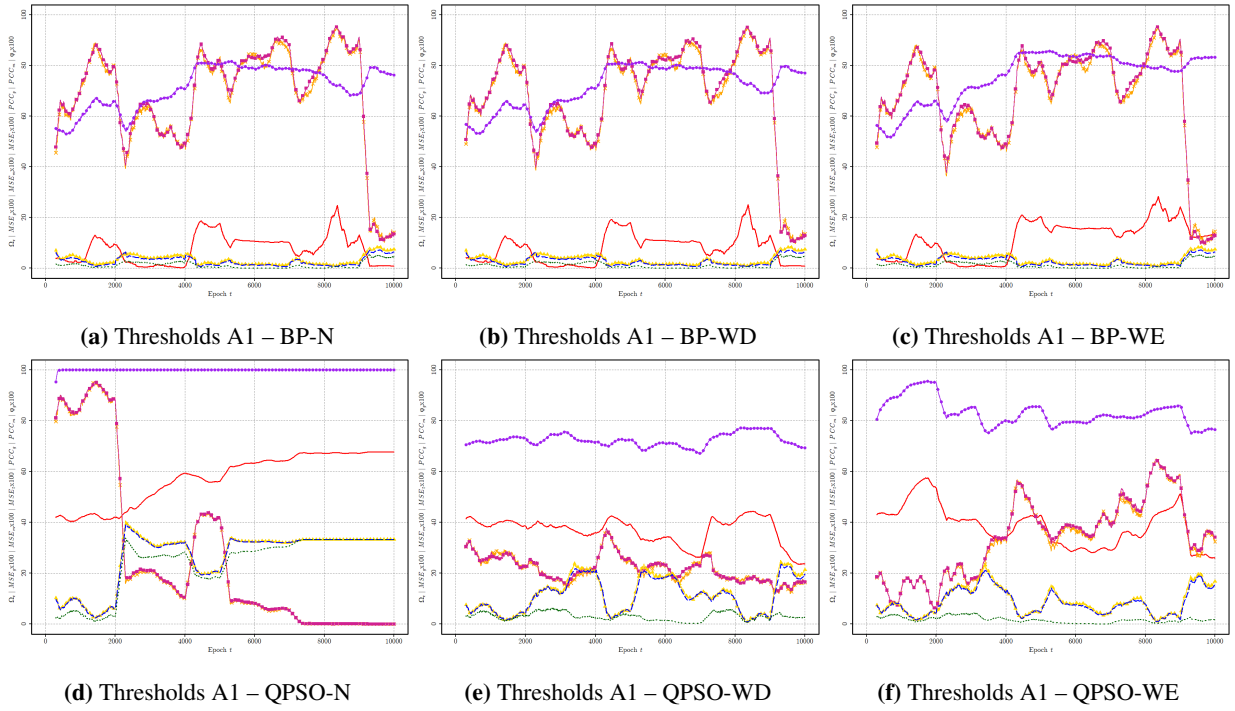


Fig. B.9: Saturation, accuracy, and complexity performance trends of the classifiers for the A1 Thresholds problems

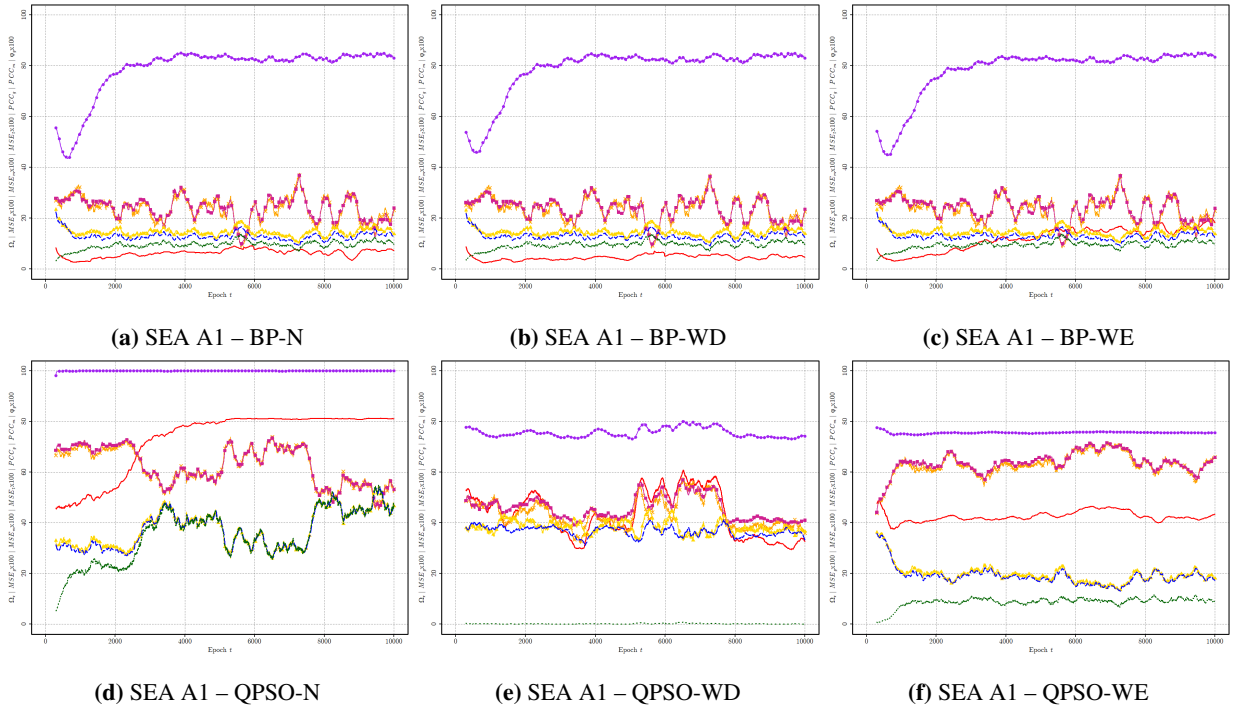


Fig. B.10: Saturation, accuracy, and complexity performance trends of the classifiers for the A1 SEA problems

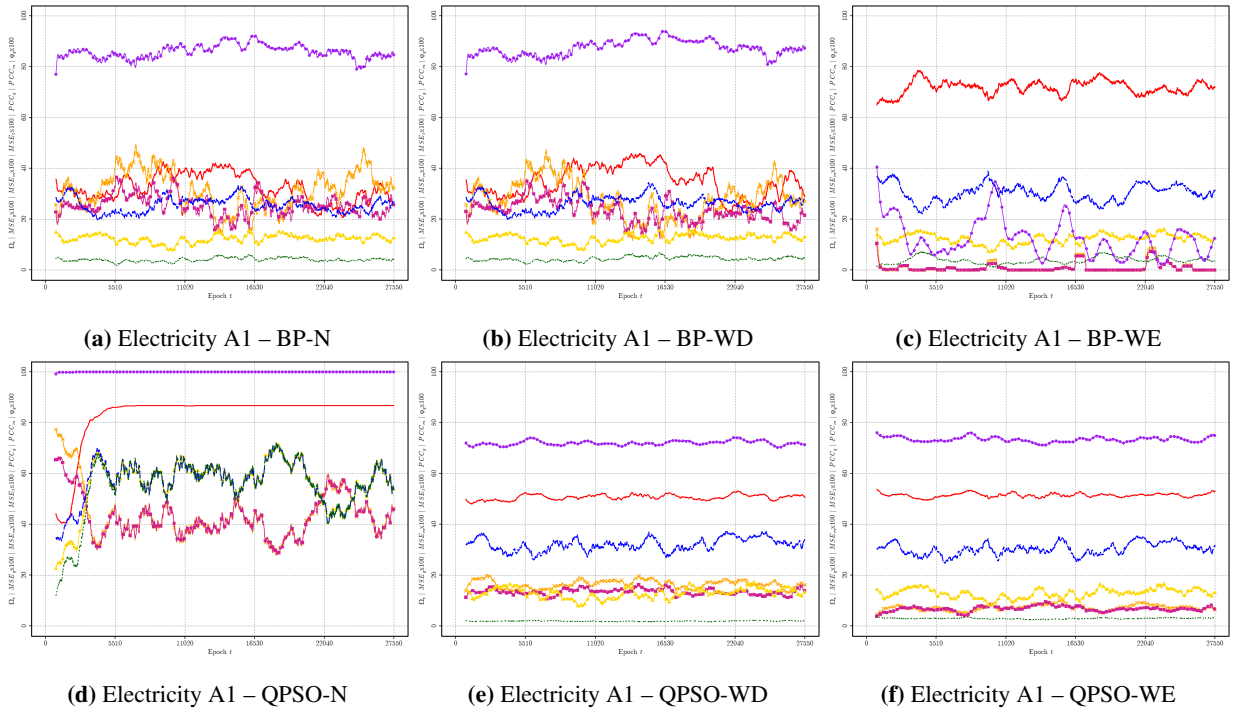


Fig. B.11: Saturation, accuracy, and complexity performance trends of the classifiers for the A1 Electricity problems

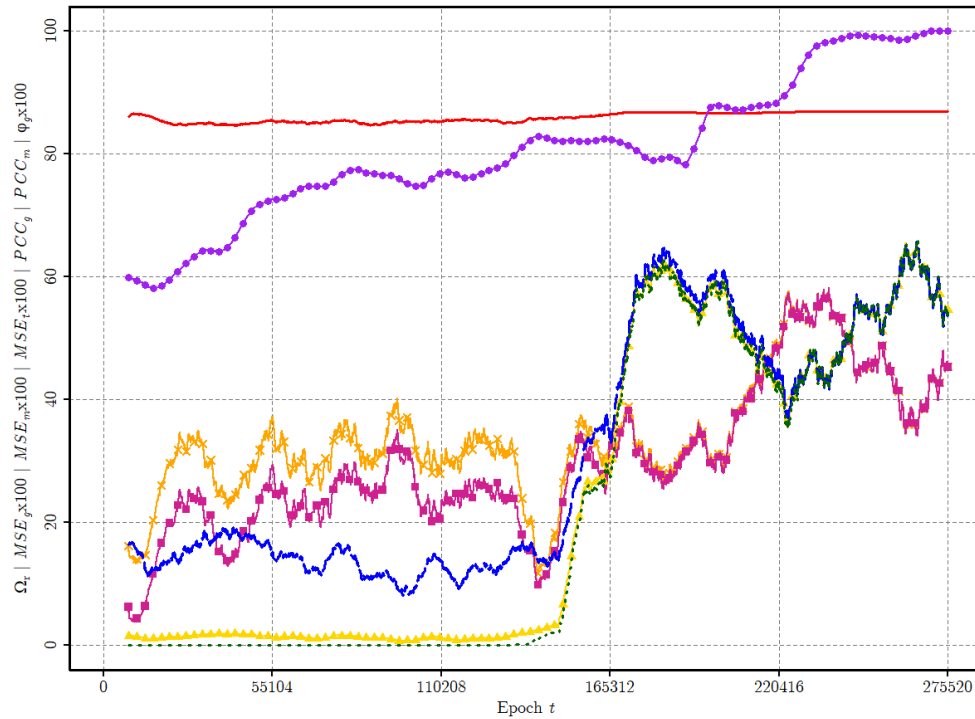


Fig. B.12: Saturation, accuracy, and complexity performance trends of the BP-WE for the A4 Electricity problem

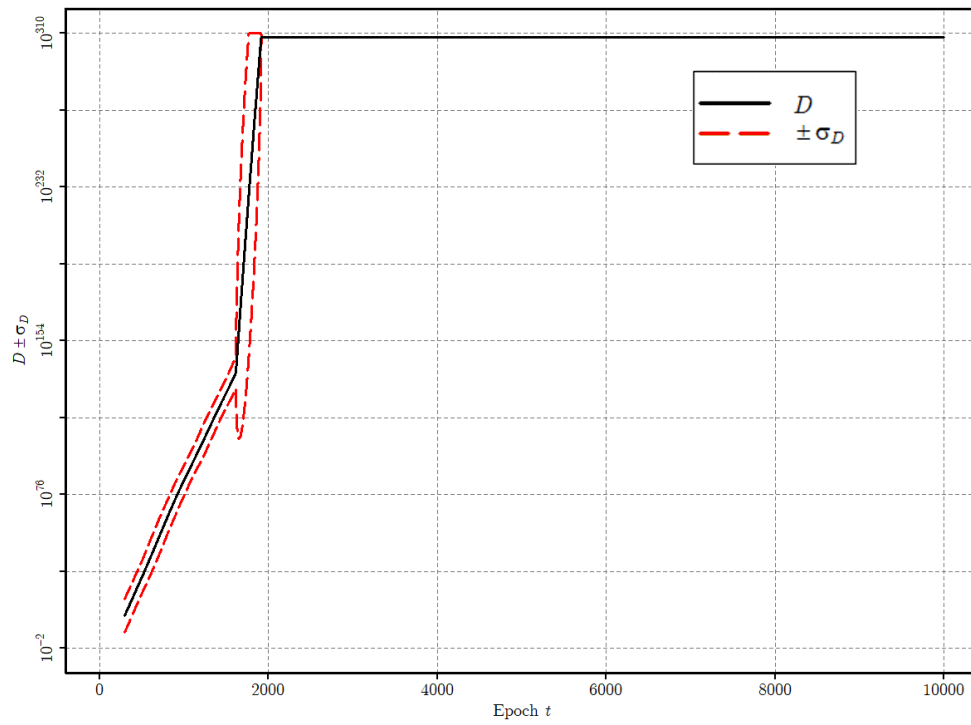


Fig. B.13: Swarm diversity trend of the QPSO-N classifier for the A1 hyperplane problem

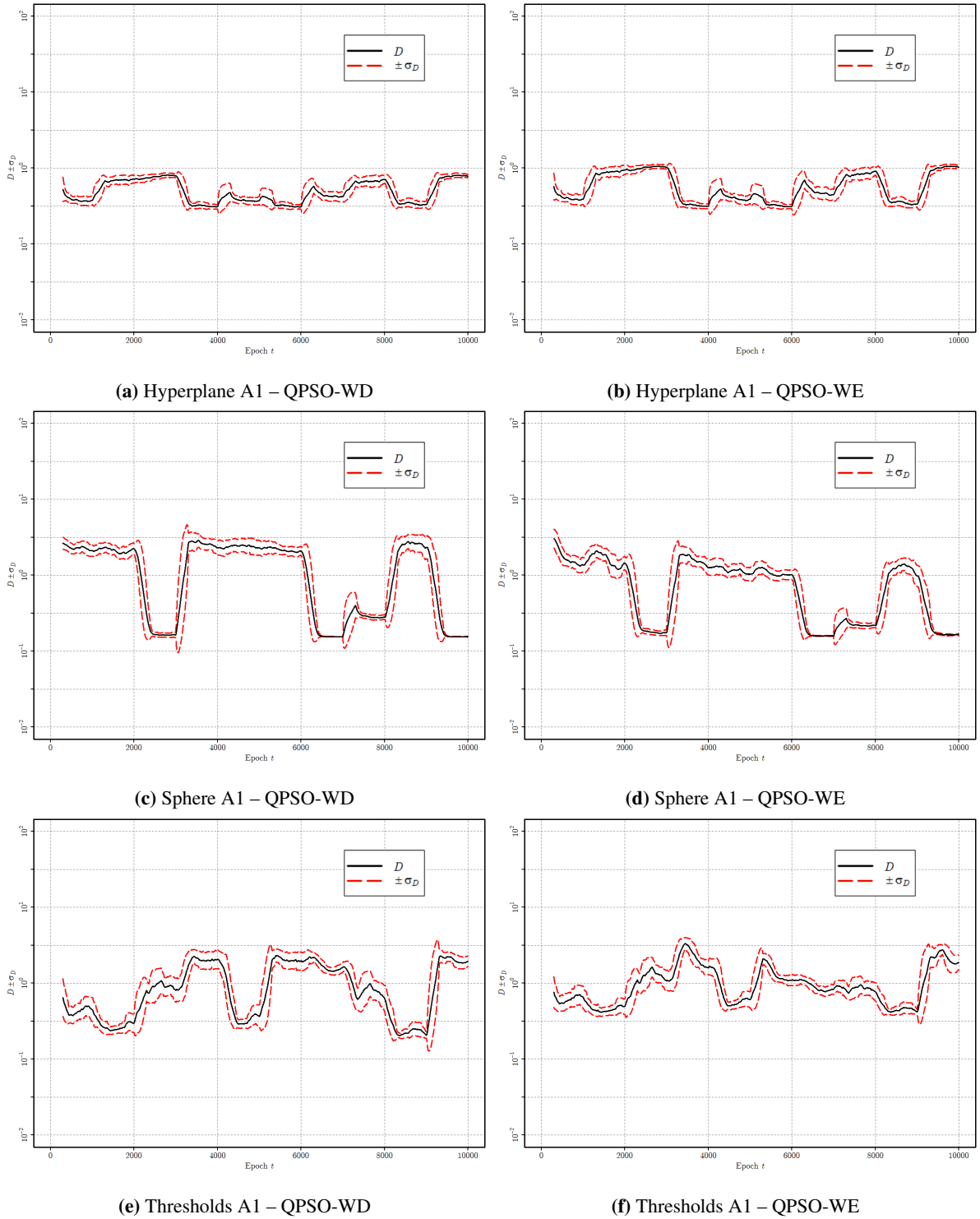


Fig. B.14: Swarm diversity trends of the QPSO-WD and QPSO-WE classifiers for the A1 hyperplane, A1 sphere and A1 thresholds problems

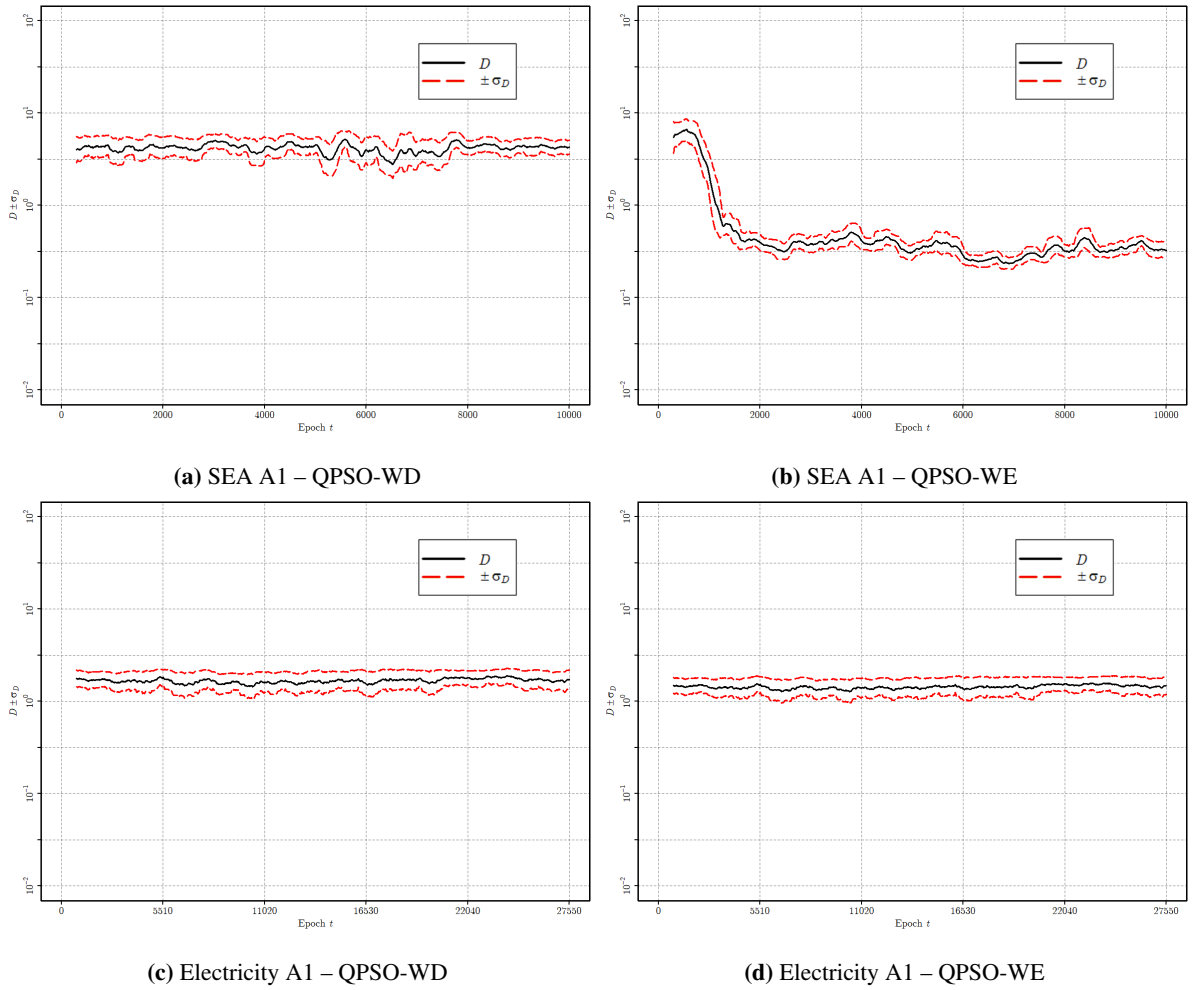


Fig. B.15: Swarm diversity trends of the QPSO-WD and QPSO-WE classifiers for the A1 SEA and A1 electricity problems

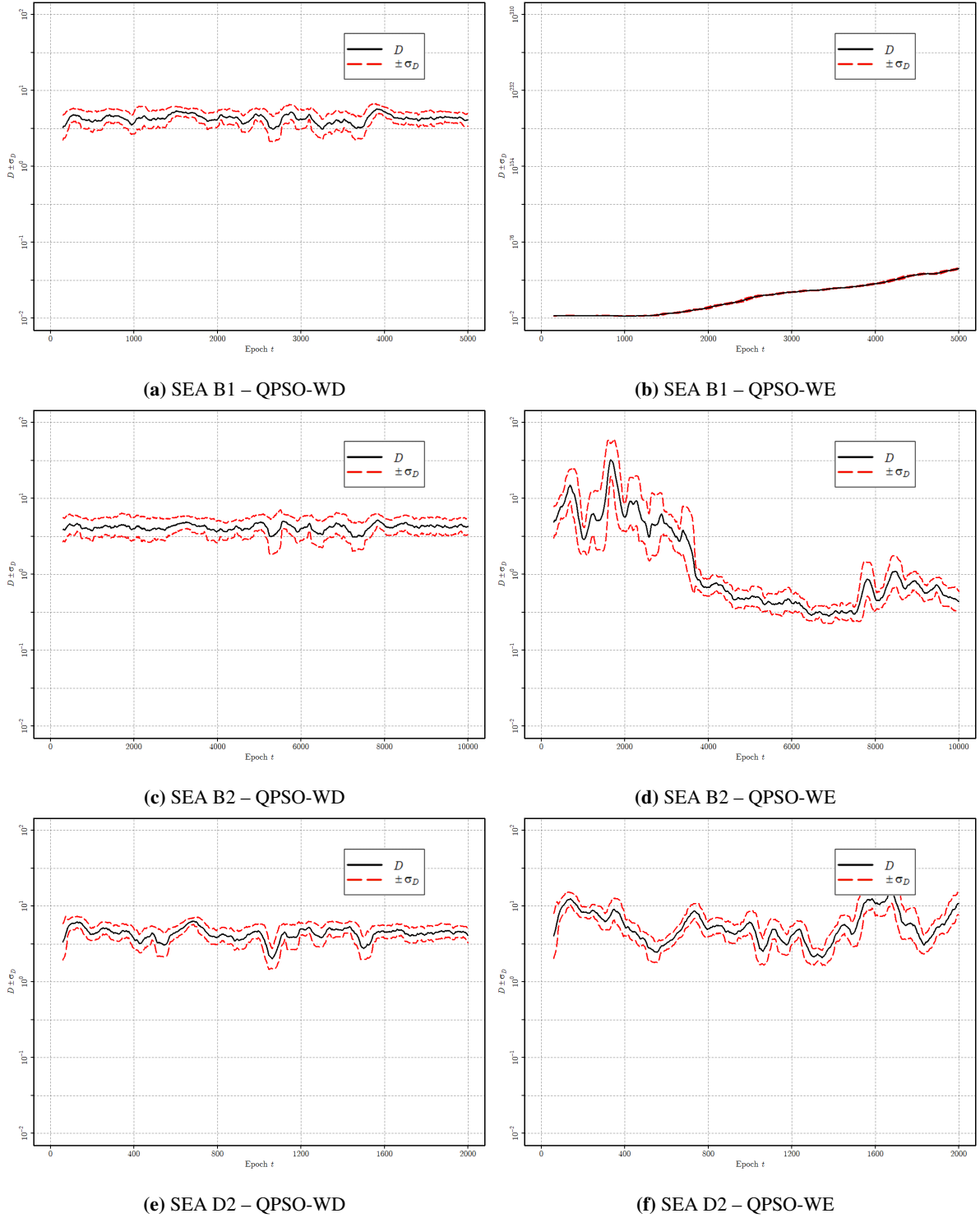


Fig. B.16: Swarm diversity trends of the QPSO-WD and QPSO-WE classifiers for selected SEA problems

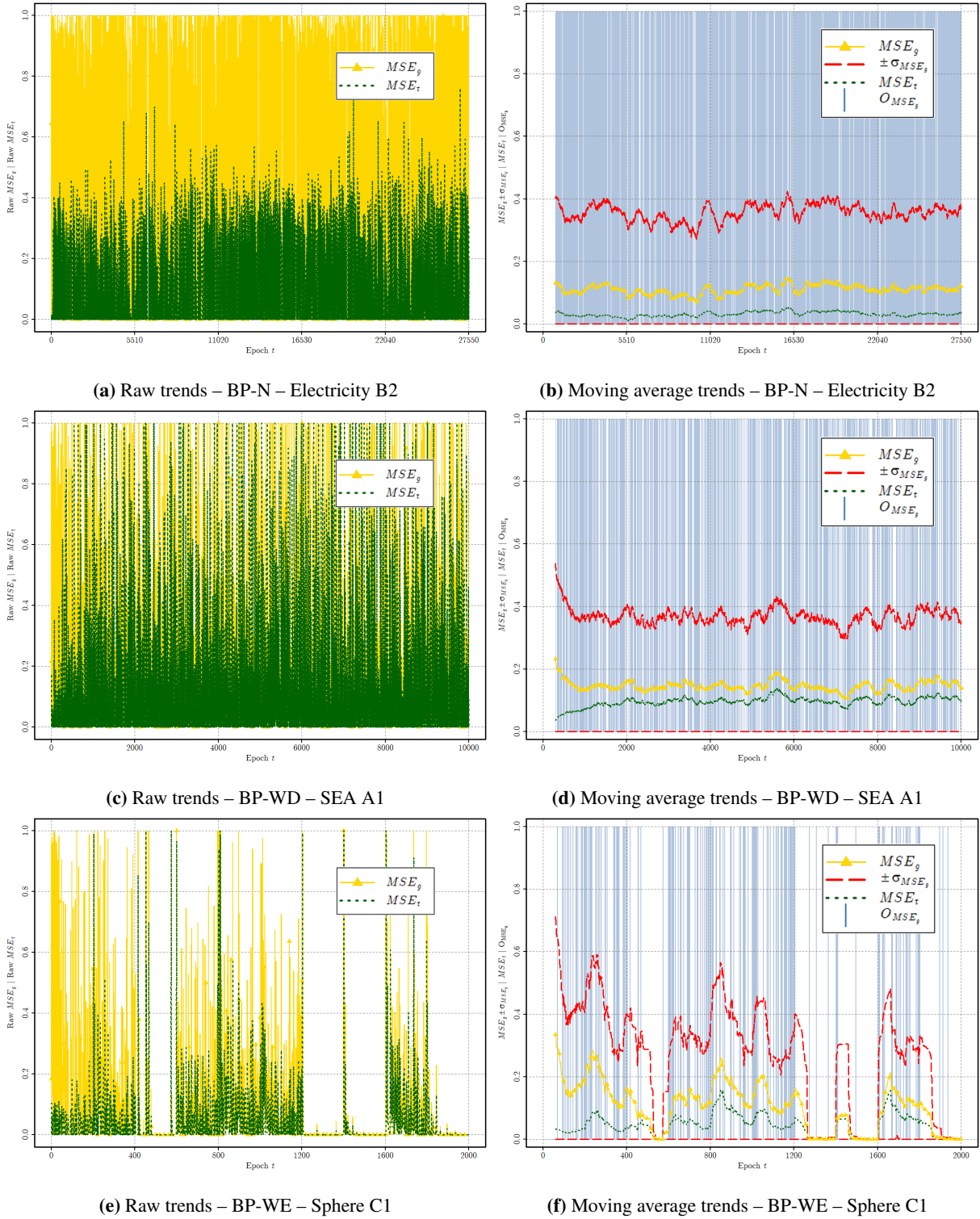


Fig. B.17: Raw MSE_g and MSE_t trends versus moving average $MSE_g \pm \sigma_{MSE_g}$ and MSE_t trends, and O_{MSE_g} for the BP classifiers

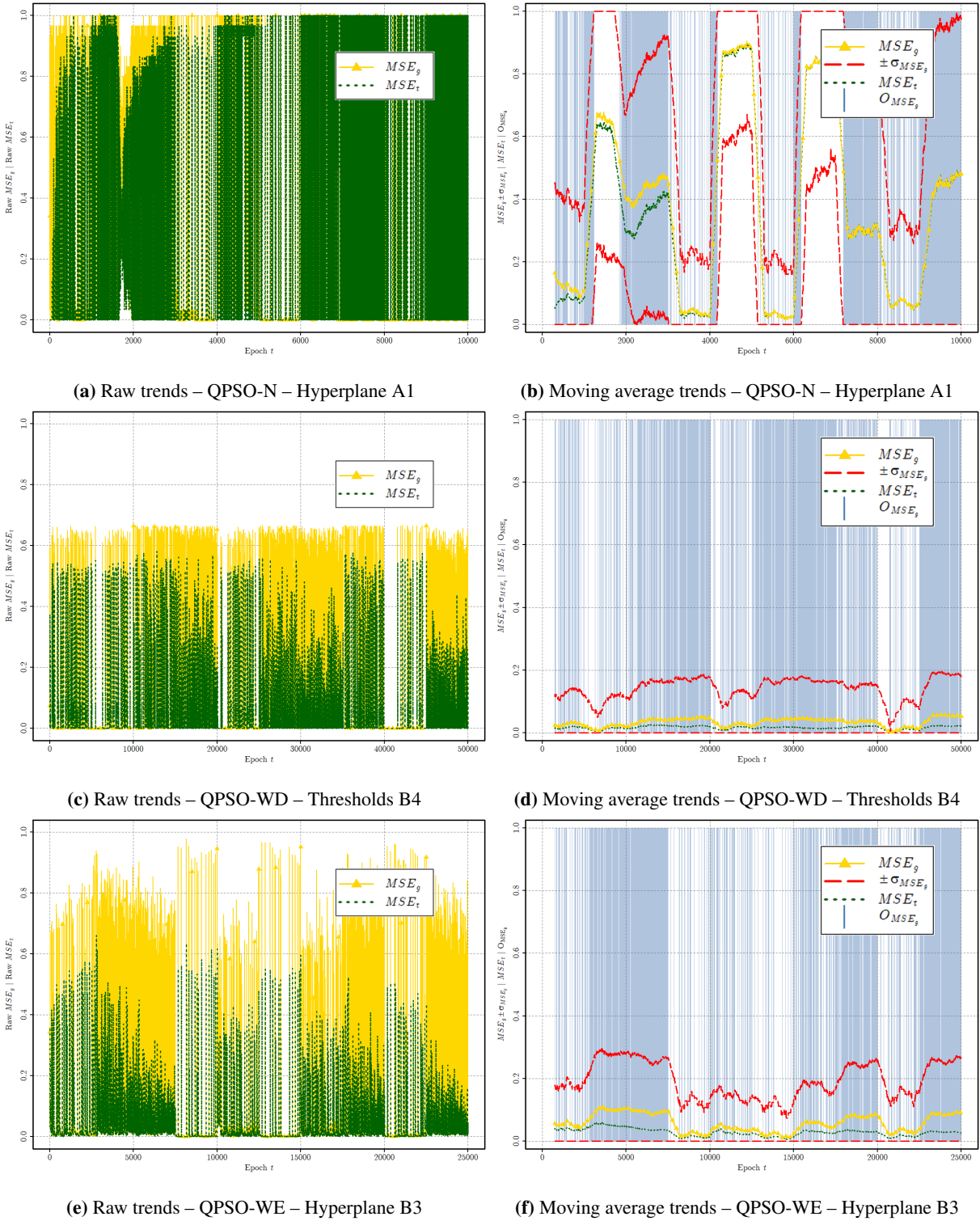


Fig. B.18: Raw MSE_g and MSE_t trends versus moving average $MSE_g \pm \sigma_{MSE_g}$ and MSE_t trends, and O_{MSE_g} for the QPSO classifiers

Appendix C. Feed forward neural network architecture diagram

Figure C.19 presents a general structural diagram of the 3-layer FFNN architecture that was used in this article. The circles represent non-bias neurons, the rounded squares represent bias neurons, the lines between neurons represent synapses, and the dashed rectangles represent the neuron layers.

In Figure C.19, z_i is the activation value of input neuron i that corresponds to the i -th element in input vector \vec{z} , h_h is the activation value of the hidden neuron h , and o_k is the activation value of the output neuron k that corresponds to the k -th element in target vector \vec{t} . Furthermore, w_{hi} is the weight of the synapse from input neuron i to hidden neuron h ; and w_{kh} is the weight of the synapse from hidden neuron h to output neuron k .

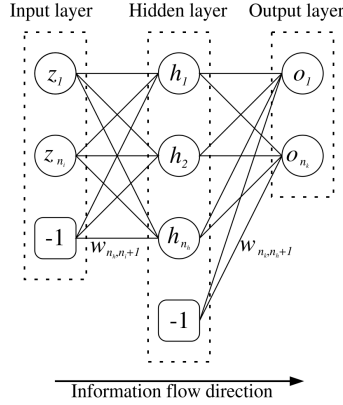


Fig. C.19: 3-layer FFNN architecture

Appendix D. Back propagation learning algorithms and equations

This appendix provides the stochastic BP algorithm and weight adjustment equations for BP-N, BP-WD and BP-WE streamed data classifiers investigated in the article.

Appendix D.1. Back propagation learning algorithm

Algorithm 1 presents the pseudo code for the BP implementation used in this article. Algorithm 1 focuses on the training of a pattern for SDCPs, and not on FFNN initialisation and training set selection. Algorithm 1 considers each training set to consist out of one pattern, because of the one-pass requirement of SDCPs.

Algorithm 1 Supervised ReLU summation 3-layer FFNN BP weights adjustment algorithm for an environment instance

```

Input:  $\eta, \alpha$ , the FFNN, pattern  $p$ 
Feed forward pattern  $p$ 's inputs into the FFNN
Calculate error signal for the output layer neurons, using Equation (D.2)
Calculate error signal for the hidden layer neurons, using Equation (D.4)
Adjust FFNN's output to hidden layer weights, using Equation (D.1)
Adjust FFNN's hidden to input layer weights, using Equation (D.3)
return FFNN

```

Each output-to-hidden layer weight, w_{kh} , was changed as follows:

$$w_{kh}(t+1) = w_{kh}(t) + \Delta w_{kh}(t) + \alpha \Delta w_{kh}(t-1) \quad (\text{D.1})$$

where $\Delta w_{kh}(t)$ is the change that weight w_{kh} experiences at epoch t , calculated as

$$\Delta w_{kh}(t) = -\eta \delta_{o_k} h_h$$

where h_h is the activation value of the h -th hidden neuron, δ_{o_k} is the error signal of the k -th output neuron, calculated as

$$\delta_{o_k} = -(t_k - o_k) f'_{ReL} \left(\sum_{h=1}^{J+1} w_{kh} h_h \right) \quad (D.2)$$

and $-(t_k - o_k)$ is the partial derivative of Equation (10) for a single pattern, and J is the number of hidden neurons excluding bias neurons in the hidden layer. f'_{ReL} is the partial derivative of $f_{ReL}(net_{o_k})$ with respect to the net input signal, net_{o_k} , which is calculated as the weighted sum of inputs into the o_k .

Each hidden-to-input layer weight, w_{hi} , was changed as follows:

$$w_{hi}(t+1) = w_{hi}(t) + \Delta w_{hi}(t) + \alpha \Delta w_{hi}(t-1) \quad (D.3)$$

where $\Delta w_{hi}(t)$ is the change that weight w_{hi} experiences at epoch t , calculated as

$$\Delta w_{hi}(t) = -\eta \delta_{h_h} z_i$$

and δ_{h_h} is the error signal of the h -th hidden neuron, calculated as

$$\delta_{h_h} = \left(\sum_{k=1}^K \delta_{o_k} w_{kh} \right) f'_{ReL} \left(\sum_{i=1}^{I+1} w_{hi} z_i \right) \quad (D.4)$$

where z_i is the value of the i -th input neuron.

Appendix D.2. Weight decay learning algorithm

To incorporate WD regularisation into Algorithm 1 the following was done: The regularisation Equation (4), with E_t as SSE_p and E_r as the WD Equation (5), was used as the new training error function, instead of SSE_p . To do so, the partial derivative of the product between the regularisation coefficient (λ_r) and the WD penalty term (Equation (5)), calculated as

$$\frac{\partial}{\partial w_n} \left[\lambda_r \frac{1}{2} \sum_{n=1}^{n_s} w_n^2 \right] = \lambda_r w_n, \quad (D.5)$$

was incorporated into the weight adjustment equations (D.1) and (D.3) as follows:

$$\Delta w_{kh}(t) = -\eta (\delta_{o_k} h_h + \lambda_r w_{kh}) \quad (D.6)$$

$$\Delta w_{hi}(t) = -\eta (\delta_{h_h} z_i + \lambda_r w_{hi}) \quad (D.7)$$

Appendix D.3. Weight elimination learning algorithm

WE regularisation was incorporated into Algorithm 1 in the same way as was done with WD regularisation in Appendix D.2. The partial derivative of the product between the regularisation coefficient (λ_r) and the WE penalty term (Equation (6)), calculated as

$$\frac{\partial}{\partial w_n} \left[\lambda_r \sum_{n=1}^{n_s} \frac{\frac{w_n^2}{w_0^2}}{1 + \frac{w_n^2}{w_0^2}} \right] = \lambda_r \frac{\partial}{\partial w_n} \left[\sum_{n=1}^{n_s} \frac{w_n^2}{w_0^2 + w_n^2} \right] \quad (D.8)$$

$$= \lambda_r \frac{2w_n w_0^2}{(w_0^2 + w_n^2)^2}, \quad (D.9)$$

was incorporated into the weight adjustment equations (D.1) and (D.3) as follows:

$$\Delta w_{kh}(t) = -\eta \left[\delta_{o_k} h_h + \lambda_r \frac{2w_{kh} w_0^2}{(w_0^2 + w_{kh}^2)^2} \right] \quad (D.10)$$

$$\Delta w_{hi}(t) = -\eta \left[\delta_{h_h} z_i + \lambda_r \frac{2w_{hi} w_0^2}{(w_0^2 + w_{hi}^2)^2} \right] \quad (D.11)$$

Appendix E. Quantum particle swarm optimisation learning algorithm

The pseudo code for the QPSO weights adjustment algorithm implementation used in this article is given by Algorithm 2. Algorithm 2 focuses only on training a swarm of particles for a pattern of an SDCP. Algorithm 2 assumes each training set only has one pattern, because of the one-pass requirement of SDCPs.

Algorithm 2 Supervised ReLU summation 3-layer FFNN QPSO weights adjustment algorithm for an environment instance

```

Input:  $\omega, c_1, c_2, r$ , distribution  $d$ , swarm of  $n_p$  particles, pattern  $p$ 
for  $i = 0$  to  $n_p$  do
  Calculate the quality of  $\vec{y}_i$ , using Equation (11) and pattern  $p$ 
end for
for  $i = 0$  to  $n_p$  do
  Set  $\vec{y}_i$  as the best personal best position in the Von Neumann neighbourhood of particle  $i$ 
  if (particle  $i$  is a quantum particle) then
    Update  $\vec{x}_i$  using Equation (3)
  else
    Update  $\vec{v}_i$  using Equation (1)
    Update  $\vec{x}_i$  using Equation (2)
  end if
end for
for  $i = 0$  to  $n_p$  do
  Calculate the quality of  $\vec{x}_i$  using Equation (11) and pattern  $p$ 
  if ( $\vec{x}_i$  is better than  $\vec{y}_i$ ) then
     $\vec{y}_i = \vec{x}_i$ 
  end if
end for
return Swarm of  $n_p$  particles

```