Supplementary Materials for “LMR-EWMA: A LASSO-based Multivariate Residual Control Chart for Monitoring Rare Health-Related Events”

Ruoyu Liu
The Second Affiliated Hospital of Xi' an Jiaotong University
Xi'an Jiaotong University
Xi'an, China
ruoyuliu@xjtu.edu.cn

Jiayin Wang
School of Computer Science and Technology
Xi'an Jiaotong University
Xi'an, China
wangjiayin@mail.xjtu.edu.cn

Xiaoyan Zhu
School of Computer Science and Technology
Xi'an Jiaotong University
Xi'an, China
zhu.xy@xjtu.edu.cn

Yuqian Liu
School of Computer Science and Technology
Xi'an Jiaotong University
Xi'an, China
yuqian.liu@xjtu.edu.cn

Yixuan Wang
Department of Biomedical Engineering
Nanjing University of Aeronautics and Astronautics
Nanjing, China
wangyixuan@nuaa.edu.cn

Xin Lai *
School of Computer Science and Technology
Xi'an Jiaotong University
Xi'an, China
laixin@xjtu.edu.cn

I. Results of Case 1 and Case 2

TABLE I. THE ARL_1S OF LMR-EWMA AND THE COMPARISON CHARTS IN CASE 1, WITH $ARL_0 = 400$.

	$\ln(1.1)$	$\ln(1.2)$	$\ln(1.3)$	$\ln(1.5)$	$\ln(1.75)$	$\ln(2)$	$\ln(2.5)$	$\ln(3)$
LMR-EWMA$_1$	**123.12**	**47.99**	**23.84**	**8.32**	**3.72**	**2.28**	**1.48**	**1.29**
LMR-EWMA$_2$	137.19	58.04	29.38	10.67	4.45	2.62	1.50	1.30
LMR-EWMA$_3$	154.56	68.70	35.97	12.87	5.33	2.93	1.60	1.30
LMR-EWMA$_4$	163.90	76.90	40.19	14.68	6.06	3.23	1.63	1.30
LMSR-EWMA$_1$	169.11	87.17	51.99	26.15	16.24	11.92	8.58	7.63
LMSR-EWMA$_2$	174.81	89.80	53.87	26.10	15.57	11.16	8.14	7.28
LMSR-EWMA$_3$	185.03	93.14	56.44	26.06	14.91	10.57	7.44	6.44
LMSR-EWMA$_4$	193.72	100.73	57.90	26.50	14.62	10.16	6.99	6.17
MSS-EWMA$_1$	175.35	88.87	52.42	26.40	15.57	11.76	9.14	8.51
MSS-EWMA$_2$	190.22	99.50	57.93	29.49	17.27	12.59	9.45	8.86
MSS-EWMA$_3$	201.19	111.84	68.99	33.14	18.70	13.59	10.07	9.24
MSS-EWMA$_4$	209.85	120.54	73.80	35.89	20.29	14.29	10.45	9.66
NM-EWMA$_1$	153.46	61.48	37.53	20.67	13.74	10.47	7.63	6.86
NM-EWMA$_2$	155.63	70.02	43.37	25.92	17.71	13.67	10.04	6.90
NM-EWMA$_3$	166.45	84.44	56.54	36.29	25.83	20.07	10.93	5.36
NM-EWMA$_4$	173.87	99.94	72.06	49.32	25.97	18.34	11.23	5.09
MSR-EWMA$_1$	263.26	134.54	74.02	34.46	20.10	14.48	10.38	9.17
MSR-EWMA$_2$	269.39	137.21	75.92	34.02	18.75	13.36	9.52	8.40
MSR-EWMA$_3$	272.60	145.35	80.02	33.75	18.13	12.33	8.58	7.70
MSR-EWMA$_4$	278.29	153.97	84.71	34.26	17.62	11.75	8.03	6.96

TABLE II. THE ARL_1S OF LMR-EWMA AND THE COMPARISON CHARTS IN CASE 2, WITH $ARL_0 = 400$.

	$\ln(1.1)$	$\ln(1.2)$	$\ln(1.3)$	$\ln(1.5)$	$\ln(1.75)$	$\ln(2)$	$\ln(2.5)$	$\ln(3)$
LMR-EWMA_1	**306.85**	**244.08**	**201.38**	**145.64**	**103.86**	**81.38**	**58.76**	**46.42**
LMR-EWMA_2	321.44	263.16	222.42	166.83	125.51	102.61	71.22	59.78
LMR-EWMA_3	328.22	283.26	241.03	183.92	149.86	115.79	89.93	76.52
LMR-EWMA_4	337.11	289.72	256.18	198.81	158.73	131.43	99.02	85.89
LMSR-EWMA_1	321.44	270.54	228.95	174.88	136.69	113.16	86.41	74.28
LMSR-EWMA_2	330.20	280.19	235.03	187.43	143.38	119.33	90.96	78.39
LMSR-EWMA_3	336.15	285.85	252.33	198.33	157.48	130.57	101.05	85.43
LMSR-EWMA_4	342.07	302.60	262.27	213.59	165.38	138.16	109.54	94.03
MSS-EWMA_1	326.97	269.21	226.71	165.76	123.71	100.62	75.82	62.86
MSS-EWMA_2	337.57	281.50	239.29	179.41	138.13	110.02	83.39	72.05
MSS-EWMA_3	348.74	287.49	249.10	194.89	150.40	125.97	96.59	82.79
MSS-EWMA_4	346.48	297.74	257.36	203.51	159.59	135.24	106.80	89.52
NM-EWMA_1	400.44	401.82	388.89	373.86	374.64	361.01	341.11	307.29
NM-EWMA_2	385.34	395.60	390.98	373.48	363.81	351.49	318.15	303.70
NM-EWMA_3	392.91	386.23	388.11	378.82	355.26	350.94	319.22	304.46
NM-EWMA_4	388.53	400.28	392.41	397.86	371.55	357.29	331.03	327.71
MSR-EWMA_1	326.97	269.21	226.71	165.76	123.71	100.62	75.82	62.86
MSR-EWMA_2	337.57	281.50	239.29	179.41	138.13	110.02	83.39	72.05
MSR-EWMA_3	348.74	287.49	249.10	194.89	150.40	125.97	96.59	82.79
MSR-EWMA_4	346.48	297.74	257.36	203.51	159.59	135.24	106.80	89.52

II. CASE STUDY OF FLIGHT DELAYS

From the official website of the US Department of Transportation (https://www.transtats.bts.gov/DL_SelectFields.asp?gnoyr_VQ=FGJ&QO_fu146_anzr=b0-gvzr), we selects flight data from 2012 to 2014 to monitor changes in the number of flight delays, focusing on daily flights from Hartsfield-Jackson Atlanta International Airport (ATL) to Orlando International Airport, Florida (MCO). The dataset comprises 21,858 records (after removing missing and outlier values), with observations based on the daily number of flights from ATL to MCO that experienced arrival delays. Using the data from 2012-2013 as Phase-I data, we train the risk adjustment model and obtain CLs to detect any increase in delay risks in 2014.

Firstly, similar to influenza patient numbers, we verify whether the number of flight delays can be treated as rare events. Applying the same three methods outlined in the main text, the results indicate that the ZIP distribution, rather than the classic Poisson distribution, better describes the number of flight delays, confirming them as rare events. These results are available from the authors.

Then, refer to Hu et al. [1], we establish a ZIP regression model for risk adjustment, using season, wind speed and air traffic control as shared covariates for both the Poisson parameter and the zero-inflated parameter. Season is represented by dummy variables; wind speed data is sourced from Weather Underground (https://www.wunderground.com/history/monthly/us); and air traffic control information is captured as the daily number of departure delays among flights from MCO.

The ARL_0, the comparison charts, and the smooting parameters employed in this case study are the same as Case Study section of the main text.

Fig.1 presents the observations and comparative experimental results based on this dataset, with red dashed lines and numerals retaining the same meaning as in the main text. The comparison results between LMR-EWMA and existing univariate residual charts is not reported here but available from the authors. Both them and Fig.1 lead to the same conclusion as the influenza dataset analysis, namely, LMR-EWMA still demonstrates the best monitoring performance.

Finally, it should be noted that in this case study, the univariate EWMA based on PR fails to detect any significant shifts, suggesting that PR loses the monitoring function on this dataset. In other words, a single residual may have a risk of failure when applied to unsuitable data or scenarios. If a research constructs control charts based only on PR in these scenarios, it will be difficult for the charts to timely detect shifts in the number of rare events. In fact, this phenomenon motivates us to focus on considering all residuals simultaneously and selecting the currently optimal residual continuously.

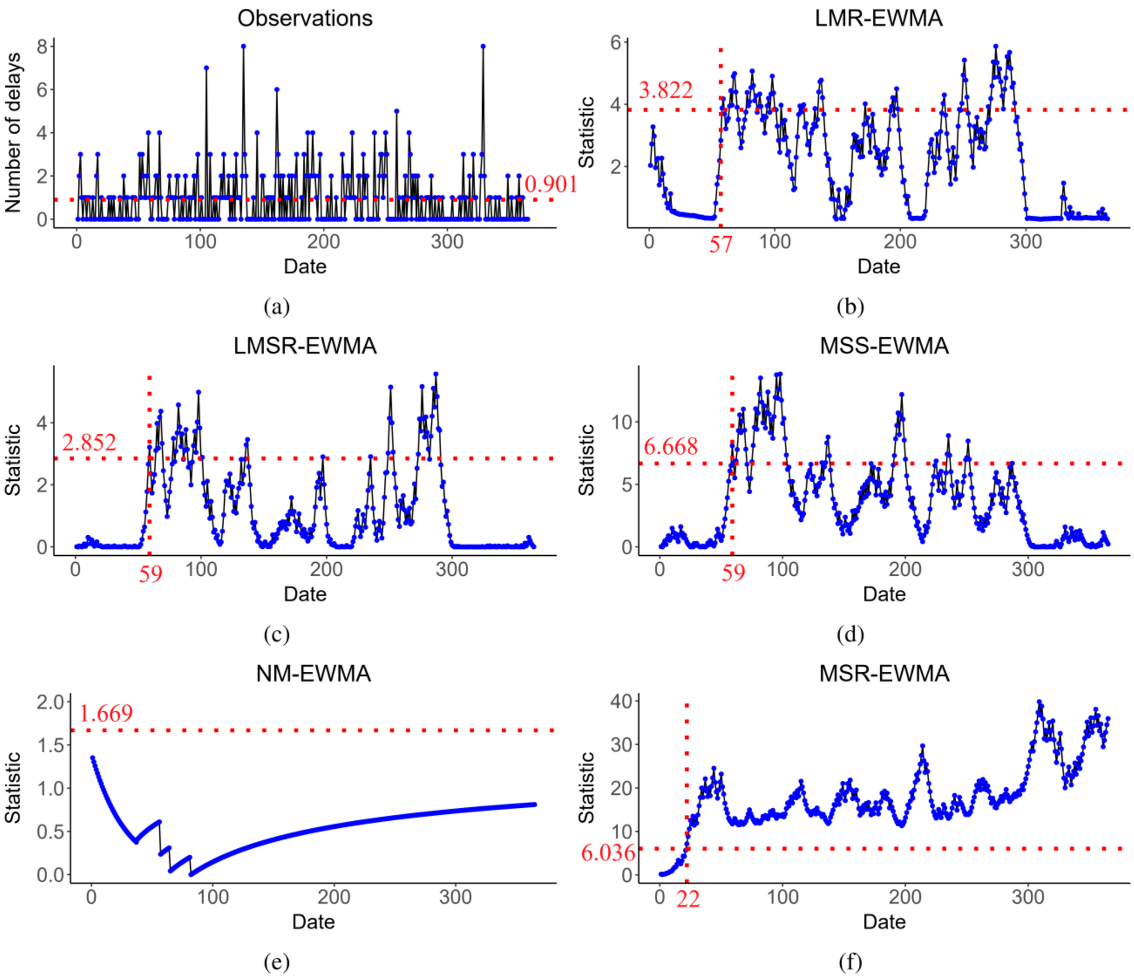

Fig. 1. The observations and chart statistics of LMR-EWMA and the comparison charts on Phase-II data. (a): The observed daily flight delay numbers in Phase-II data, with the red dashed line representing the mean of the observations. (b)-(f): The chart statistic curves for each chart, with horizontal and vertical red dashed lines and numbers representing hs and run lengths at the first alarm, respectively.

REFERENCES

[1] Q. Hu, L. Liu. Weighted score test based EWMA control charts for Zero-Inflated Poisson models[J]. Computers & Industrial Engineering, 2020:106966.