Supplement of

Brief communication: SWM: Stochastic Weather Model for precipitation-related hazard assessments

Melody G. Whitehead, Mark S. Bebbington

Correspondence to: Melody G. Whitehead (m.whitehead@massey.ac.nz)

Overview

This supplement provides the complete set of results from the statistical analyses on both exemplar samples: *Sample 1* and *Sample 2* each of which represent 95 simulated datasets. These are compared to the single real dataset for the Tarawera/Rangitāiki catchment: hourly rainfall data across a 11x14 grid of longitude: {176° E, 176.1° E, …, 177.0° E} and latitude: {37.8° S, 37.9° S, …, 39.1° S}, for 40 years (1981 – 2020) downloaded from ERA5-land.

1 Monthly means and variance

The Shapiro-Wilks normality test (Royston, 1982) was run using the *shapiro.test()* function from *base* R. A significance level of 0.05 was applied, i.e., any test result with $p < 0.05$ was deemed to have failed the normality test (Supp. Tab. 1).

The Levene test for equal variance was run using the *leveneTest()* function from *car* (Fox and Weisburg, 2019). A significance level of 0.05 was applied, i.e., any test result with $p < 0.05$ was deemed to have failed the equality of variance test (Supp. Tab. 2).

Student's t-test for equality of means (Student, 1908) was run using the *t.test()* function from *base* R. A significance level of 0.05 was applied, i.e., any test result with $p < 0.05$ was deemed to have failed the normality test (Supp. Tab. 3).

Supplementary Table 1: Shapiro-Wilks normality test failure rate (%) by month on mean monthly rainfall

Month	Real		Sample 1	Sample 2
January	p = 0.0004647	1/1 = 100 %	75/95 = 79 %	81 / 95 = 85 %
February	p = 0.001508	1/1 = 100 %	76/95 = 80 %	69 / 95 = 73 %
March	p = 0.001861	1/1 = 100 %	49/95 = 52 %	56 / 95 = 59 %
April	p = 0.009444	1/1 = 100 %	59/95 = 62 %	67 / 95 = 71 %
May	p = 0.003474	1/1 = 100 %	57/95 = 60 %	59 / 95 = 62 %
June	p = 0.06495	0/1 = 0 %	31/95 = 33 %	35 / 95 = 37 %
July	p = 0.017	1/1 = 100 %	43/95 = 45 %	42 / 95 = 44 %
August	p = 0.0947	0/1 = 0 %	40/95 = 42 %	34 / 95 = 36 %
September	p = 0.5795	0/1 = 0 %	26/95 = 27 %	24 / 95 = 25 %
October	p = 0.404	0/1 = 0 %	42/95 = 44 %	40 / 95 = 42 %
November	p = 0.0003	1/1 = 100 %	38/95 = 40 %	46 / 95 = 48 %
December	p = 0.434	0/1 = 0 %	58/95 = 61 %	35 / 95 = 37 %
Overall	**7 / 12 = 58.3 %**		**594 / 1140 = 52.1 %**	**588 / 1140 = 51.6**

Supplementary Table 2: Levene variance equality test failure rate (%) by month on mean monthly rainfall pairs (real:sample)

Month	Real : Sample 1	Real : Sample 2
January	0 / 95 = 0 %	1 / 95 = 1 %
February	7 / 95 = 7 %	12 / 95 = 12.6 %
March	0 / 95 = 0 %	0 / 95 = 0 %
April	1 / 95 = 1 %	1 / 95 = 1 %
May	3 / 95 = 3 %	1 / 95 = 1 %
June	26 / 95 = 27 %	12 / 95 = 12.6 %
July	0 / 95 = 0 %	0 / 95 = 0 %
August	0 / 95 = 0 %	1 / 95 = 1 %
September	10 / 95 = 10.5 %	2 / 95 = 2 %
October	1 / 95 = 1 %	10 / 95 = 10.5 %
November	0 / 95 = 0 %	0 / 95 = 0 %
December	3 / 95 = 3 %	3 / 95 = 3 %
Overall	**51 / 1140 = 4.5 %**	**43 / 1140 = 3.8 %**

Supplementary Table 3: Student's t test failure rate (%) by month on mean monthly rainfall pairs (real:sample)

Month	Real : Sample 1	Real : Sample 2
January	0 / 95 = 0 %	0 / 95 = 0 %
February	1 / 95 = 1 %	0 / 95 = 0 %
March	0 / 95 = 0 %	0 / 95 = 0 %
April	1 / 95 = 1 %	0 / 95 = 0 %
May	0 / 95 = 0 %	0 / 95 = 0 %
June	3 / 95 = 3 %	3 / 95 = 3 %
July	0 / 95 = 0 %	0 / 95 = 0 %
August	0 / 95 = 0 %	0 / 95 = 0 %
September	0 / 95 = 0 %	0 / 95 = 0 %
October	0 / 95 = 0 %	0 / 95 = 0 %
November	2 / 95 = 2 %	0 / 95 = 0 %
December	0 / 95 = 0 %	0 / 95 = 0 %
Overall	**8 / 1140 = 0.7 %**	**3 / 1140 = 0.3 %**

2 Significance of month and source for rainfall prediction

Linear models were built for each real:sample set using the *lm()* function in base R with both month and source as a factor, with m1 allowing for an interaction term (between month and source), and m2 not. Accompanying files: check2_Tukey_Sample1.txt and check2_Tukey_Sample2.txt provide all outputs from check 2 for each sample. While the *summary()* function for each model does provide an estimate of whether a model coefficient is statistically significant, these p-values are unreliable for pair-wise comparisons because the probability of false detection is over inflated (the family-wise error rate). Thus, Tukey's Honest Significant Difference function is used instead, this was applied using the *TukeyHSD()* function from *base* R. In all cases, the p-value was high (> 0.05), indicating that source is not a statistically significant factor in the prediction of rainfall data.

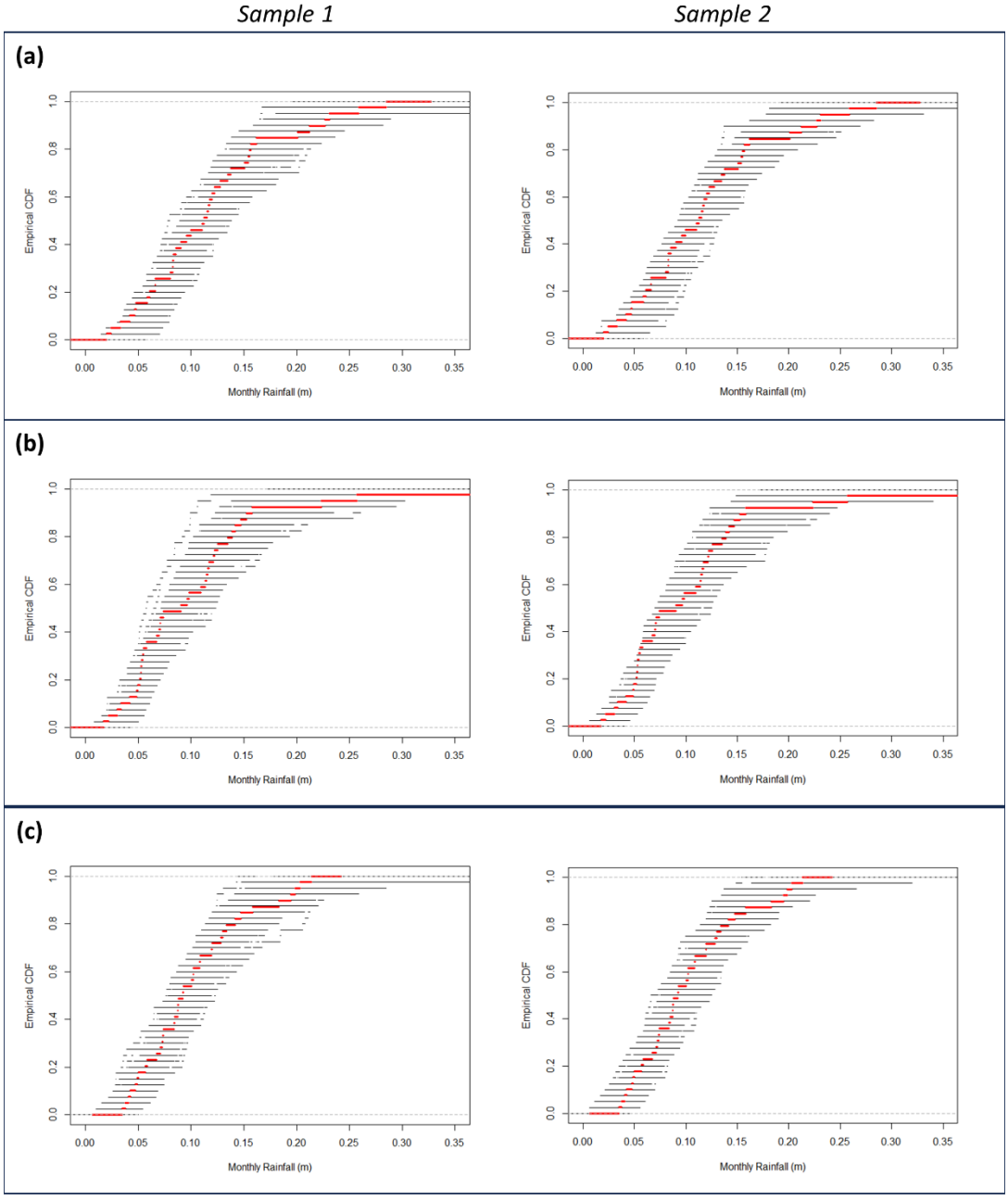
Sample 1, model 1: *TukeyHSD()* for whether source is a statistically significant factor, p = 0.8434212
Sample 1, model 2: *TukeyHSD()* for whether source is a statistically significant factor, p = 0.8434136
Sample 2, model 1: *TukeyHSD()* for whether source is a statistically significant factor, p = 0.9785945
Sample 2, model 2: *TukeyHSD()* for whether source is a statistically significant factor, p = 0.9785936

3 Distribution of monthly rainfall totals

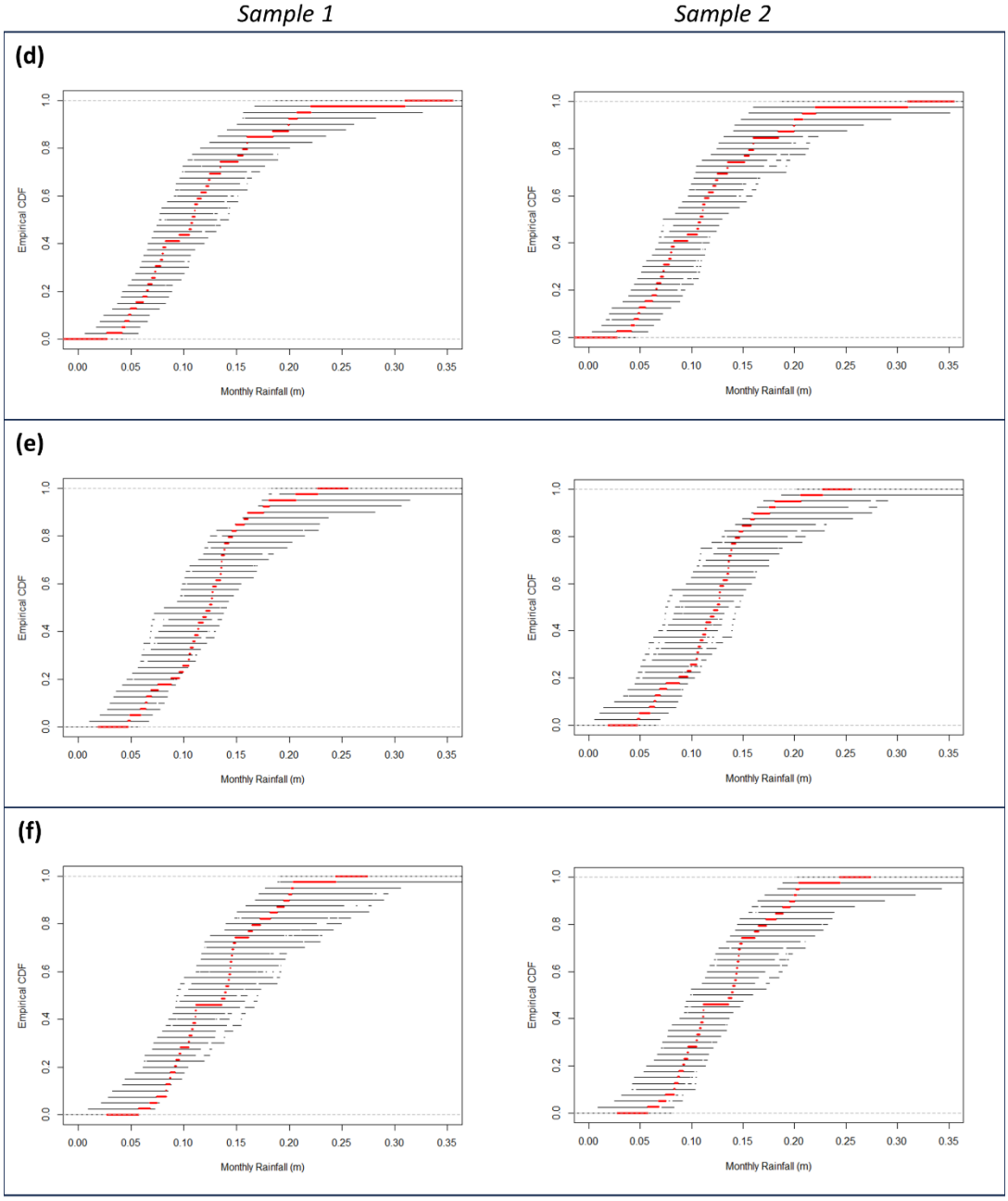
For each of the 95 simulated data in each sample, empirical Cumulative Distribution functions (eCDFs) were built using the *ecdf()* function in *base* R. These were then plotted and overlain by the real eCDF to look for departures, i.e., any locations where the real data fell outside of the envelope drawn by the simulated data. Results for *Sample 1* and *Sample 2* are provided in parallel (by month) as Supplementary Figure 1.

4 Temporal trends on daily and monthly timescales

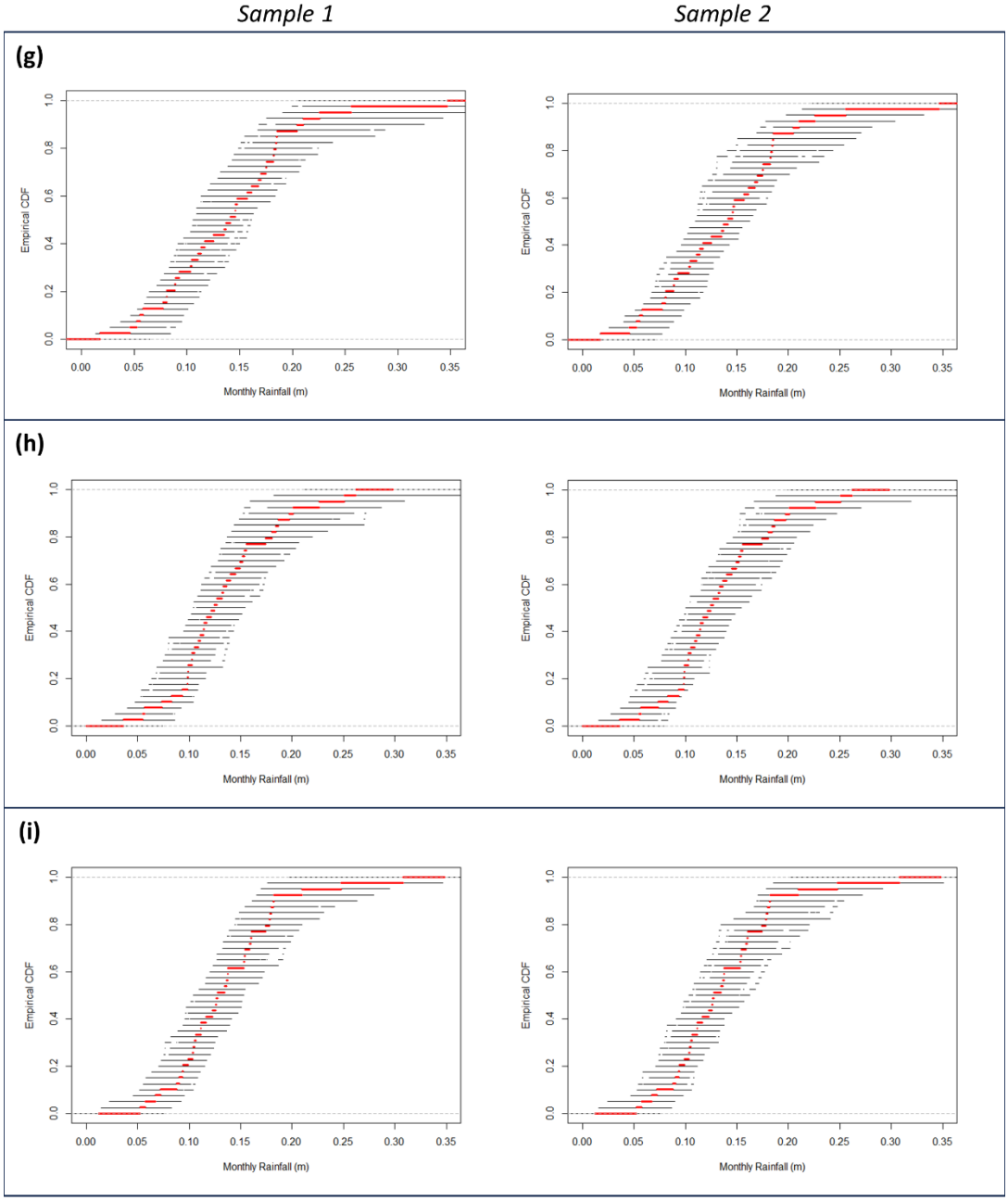
For each of the 95 simulated data in each sample, autocorrelation functions (Venables and Ripley, 2002) were built using the ac*f()* function in *base* R. These were then plotted and overlain by the real autocorrelation function to look for departures, i.e., any locations where the real data fell outside of the envelope drawn by the simulated data. Results for *Sample 1* and *Sample 2* are provided in parallel (by month) as Supplementary Figure 2.



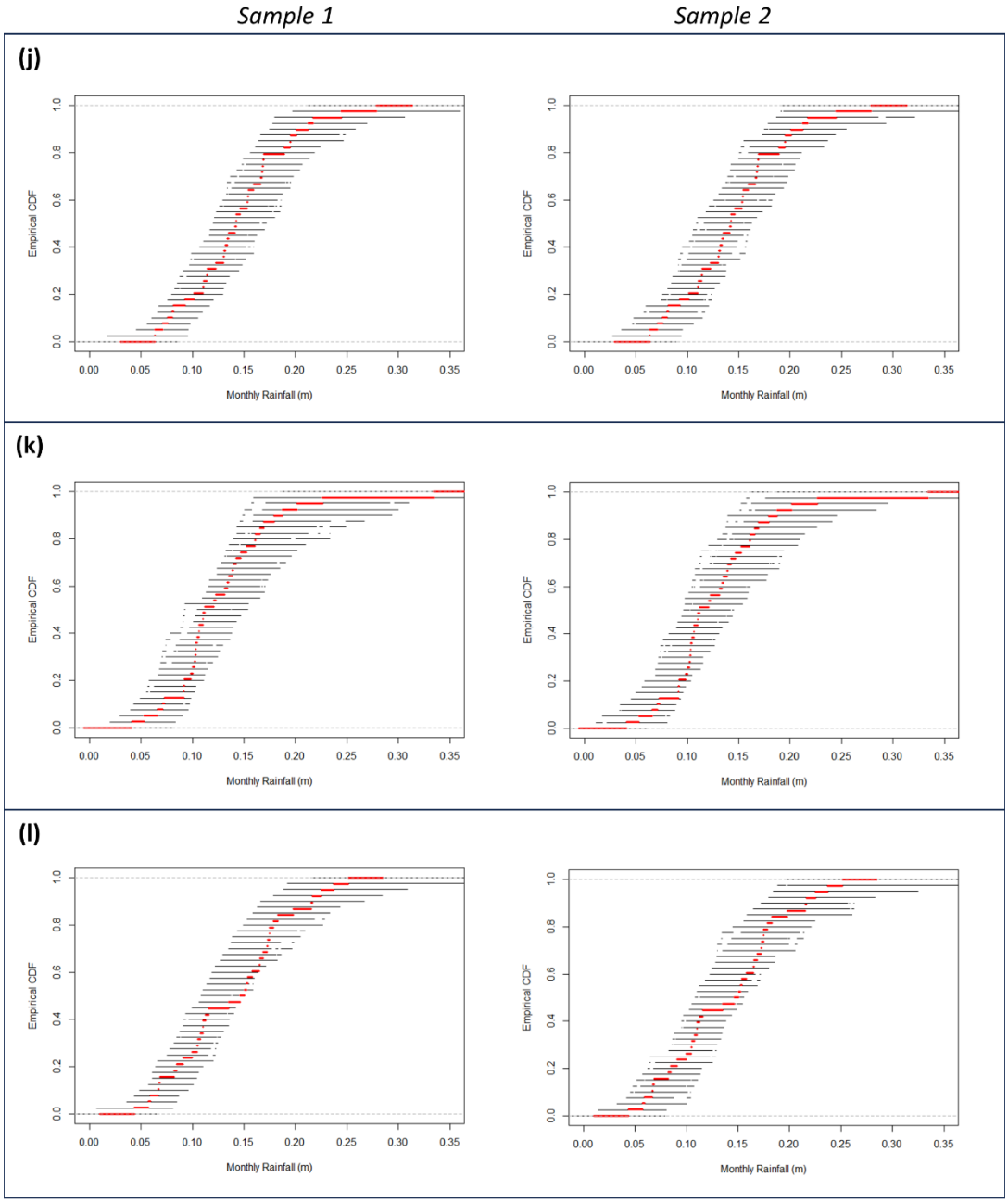
Supplementary Figure 1: Empirical Cumulative Distribution Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (a) January, (b) February, and (c) March.



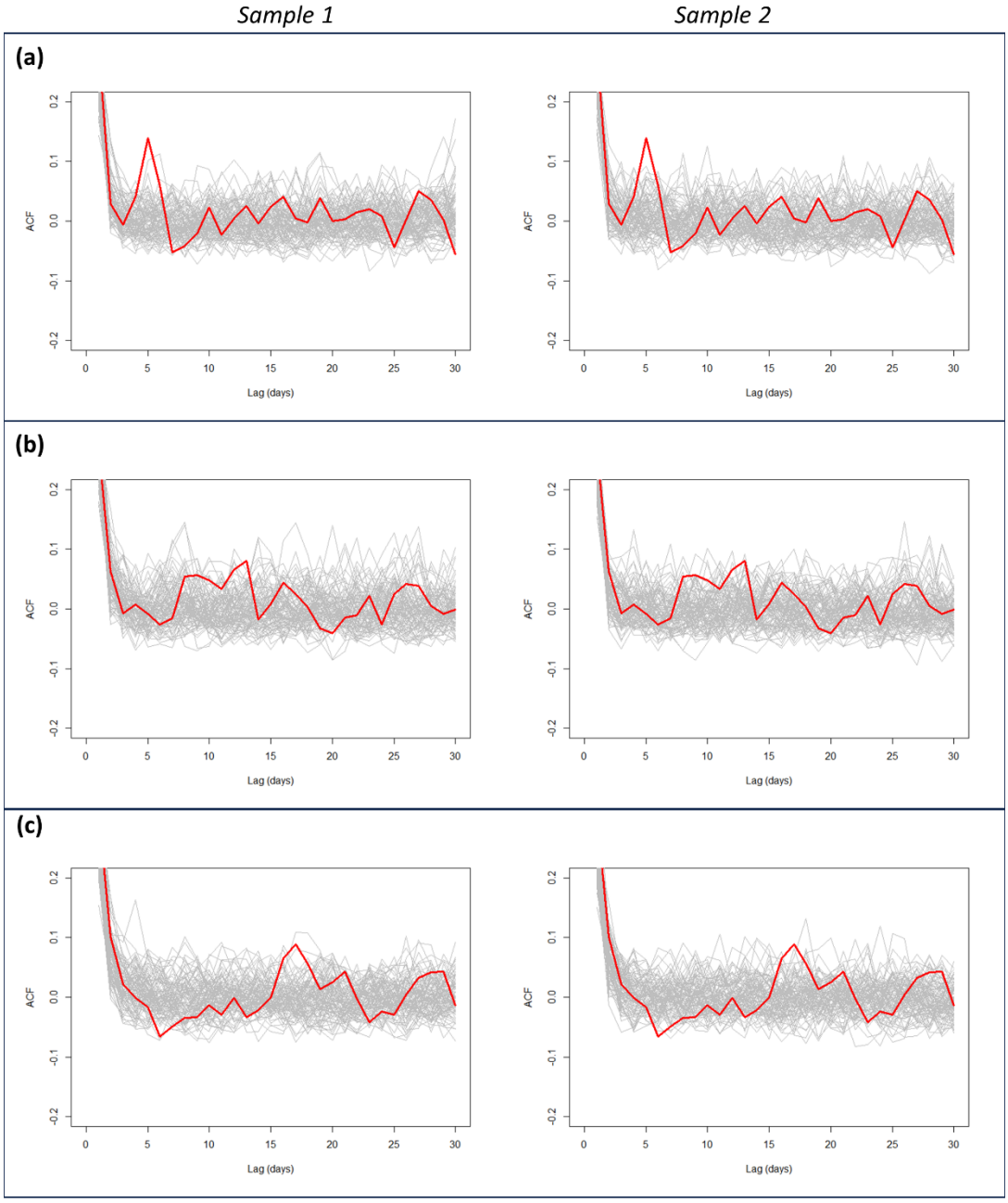
Supplementary Figure 1: Empirical Cumulative Distribution Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (d) April, (e) May, and (f) June.



Supplementary Figure 1: Empirical Cumulative Distribution Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (g) July, (h) August, and (i) September.

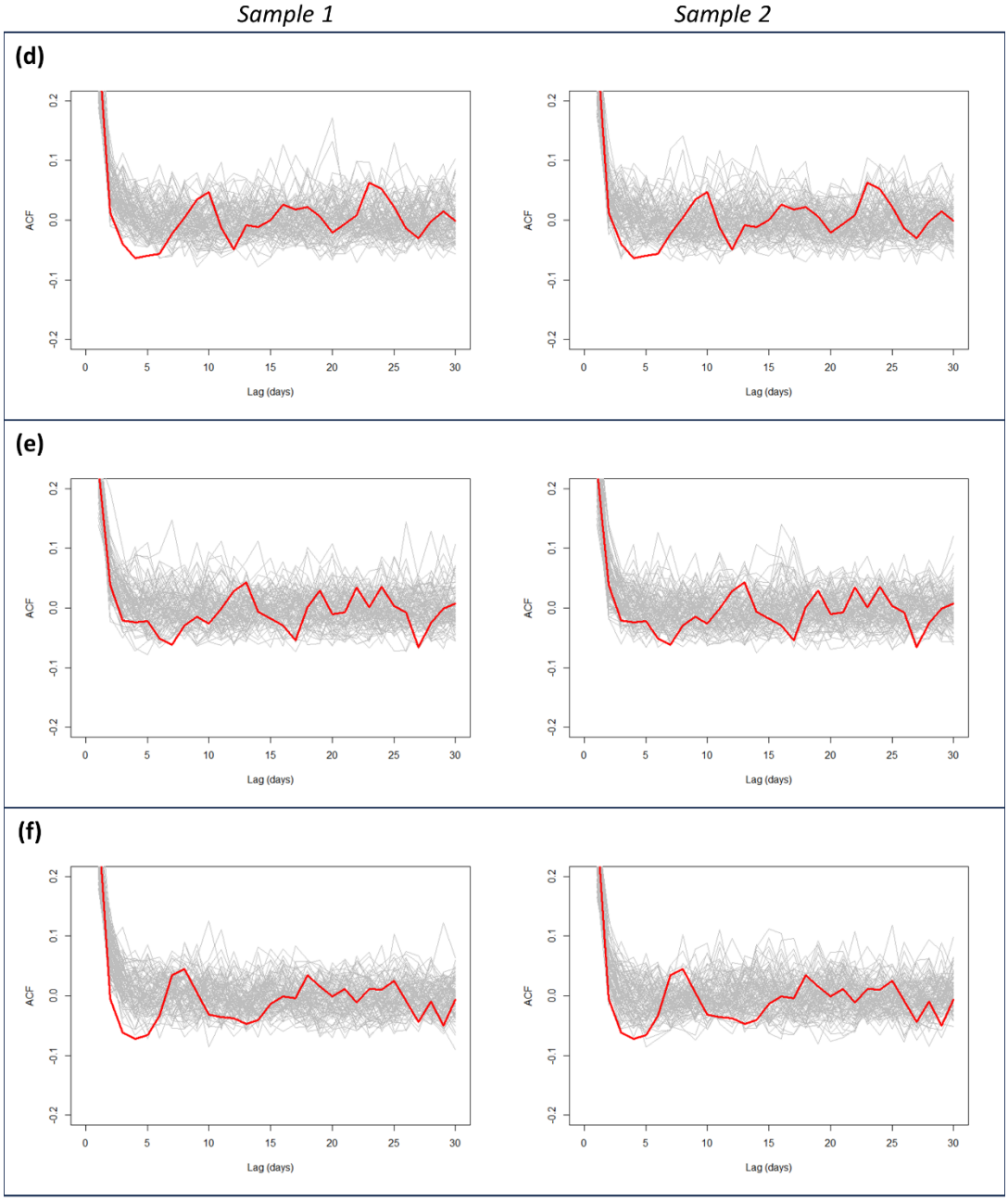


Supplementary Figure 1: Empirical Cumulative Distribution Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (j) October, (k) November, and (l) December.

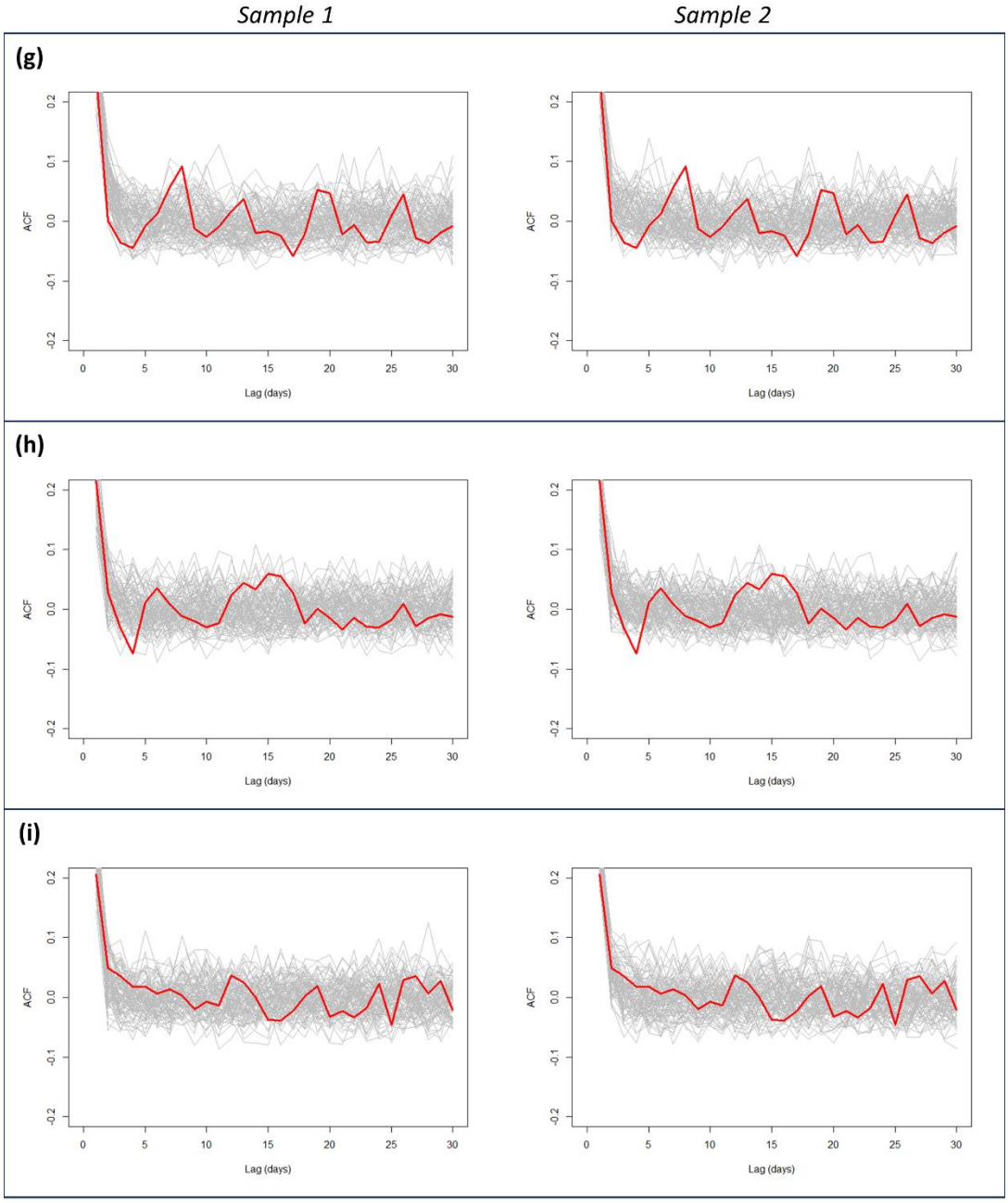


Supplementary Figure 2: Autocorrelation Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (a) January, (b) February, and (c) March.

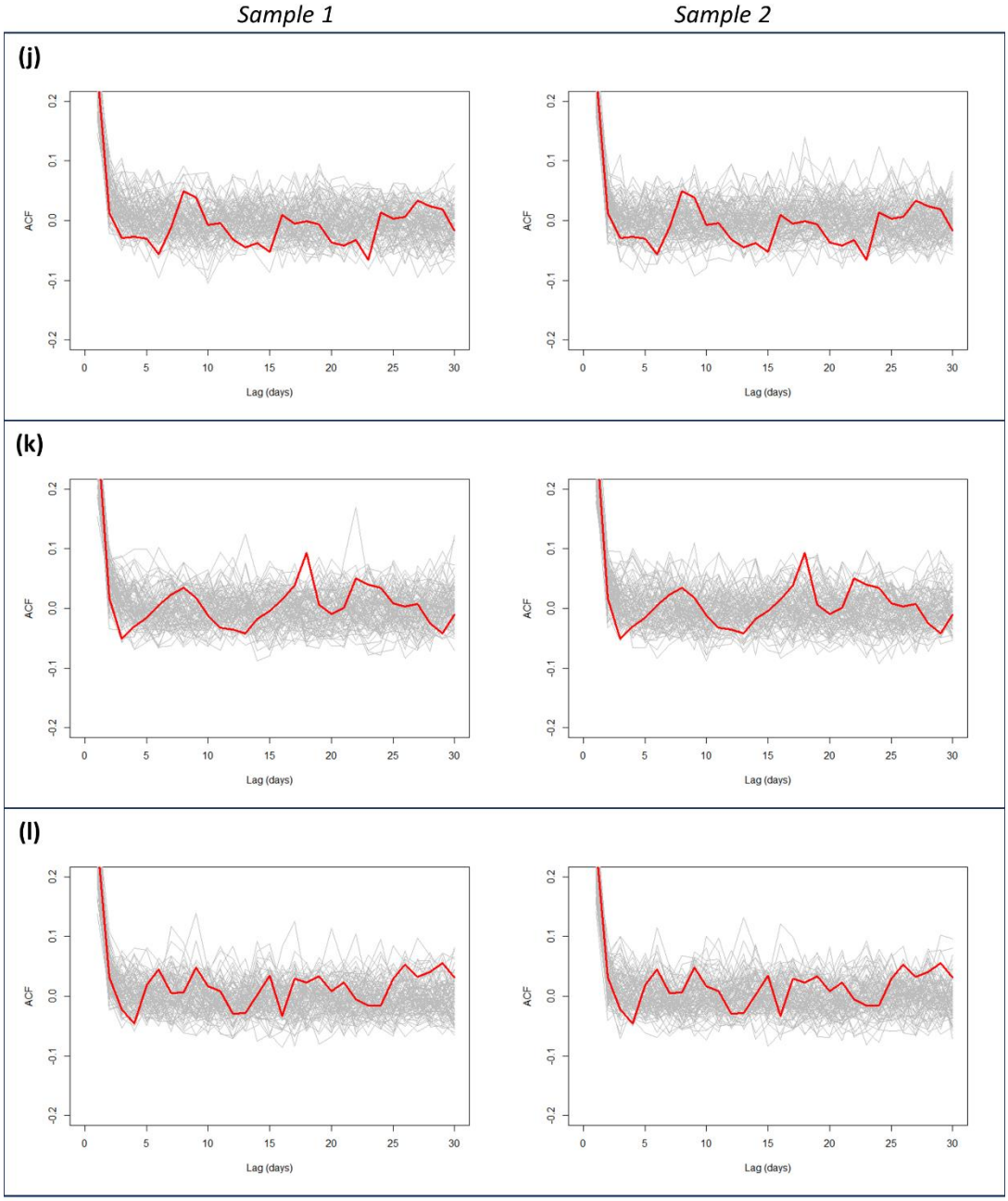
75



Supplementary Figure 2: Autocorrelation Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (d) April, (e) May, and (f) June.



Supplementary Figure 2: Autocorrelation Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (g) July, (h) August, and (i) September.



Supplementary Figure 2: Autocorrelation Functions for simulated (grey) and real (red) data for *Sample 1* and *Sample 2*, for (j) October, (k) November, and (l) December.

References

Fox. J. and Weisberg S.: An R Companion to Applied Regression, Third edition, Sage Publications, Thousand Oaks CA, USA, 576 pp., https://socialsciences.mcmaster.ca/jfox/Books/Companion/, 2019.

Royston, P.: An extension of Shapiro and Wilk's W test for normality to large samples. Appl. Stat., 31, 115–124, https://doi.org/10.2307/2347973, 1982.

Student.: The probable error of a mean, Biometrika, 1–25, 1908.

Venables, W.N. and Ripley, B. D.: Modern Applied Statistics with S. Fourth Edition, Springer Series in Statistics and Computing, Springer-Verlag, New York, NY, 498 pp., https://doi.org/10.1007/978-0-387-21706-2, 2002.