Department of Statistics
STATS 326: Applied Time Series
First Semester, 2019
Test 2 – Answer Guide
Total Marks = 100

1. The Monthly Artic Sea Ice series has a slight decreasing trend and a reasonably constant seasonal component. The seasonal peak appears to be in March with the seasonal trough in September. The seasonal component is slightly larger in 2007, 2008 and 2012. There will most likely be an autocorrelation pattern in the Residual Series from a model of these data.

(5 marks)

2. The plot of the Residual Series shows reasonably constant variation but there appears to be slight clustering. The acf of the Residual Series shows 2 significant lags (lag 1 and lag 11) which are both positive and too large to ignore. The 5-number summary of the Residual Series has similar absolute values for the min and max and also for the 1st and 3rd quartiles with a median close to 0 indicating the residuals are normally distributed. The autocorrelation in the Residual Series means the assumptions are not satisfied.

(10 marks)

3. Jan 18:
4.9029817 – 0.0014629 * 217 + 0.7387206 * 11.74
= 13.25811224 (= 13.26 millions of square kilometres)

Feb 18:
4.9029817 – 0.0014629 * 218 – 0.2676245 * 1 + 0.7387206 * 13.25811224
= 14.11048563 (= 14.11 millions of square kilometres)

Mar 18:
4.9029817 – 0.0014629 * 219 – 0.6930475 * 1 + 0.7387206 * 14.11048563
= 14.31326551 (= 14.31 millions of square kilometres)

RMSEP
= sqrt(1/3*[(13.06 – 13.25811224)^2 + (13.95 – 14.11048563)^2 +
(14.30 – 14.31326551)^2])
= 0.1473998 (= 0.15 millions of square kilometres)

(15 marks)

4. Exactly as given in Q3.

(5 marks)

5. RMSEP
= sqrt(1/3*[(13.06 – 13.21336)^2 + (13.95 – 14.08475)^2 + (14.30 – 14.3152)^2])
= 0.1181918 (= 0.12 millions of square kilometres)

(10 marks)

6. The model `SARIMA.Ice.fit` differences the trend and seasonal component. The resulting series appears to be stationary with 1 large positive residual. The plot of the acf shows significant negative autocorrelations at lags 2 and 12.

(5 marks)

7. $(1-\rho_1\mathrm{B})(1-\mathrm{B})(1-\mathrm{B}^{12})y_t = (1+\alpha_1\mathrm{B}+\alpha_2\mathrm{B}^2)(1+A_1\mathrm{B}^{12})\varepsilon_t$
=> $(1-\rho_1\mathrm{B})(1-\mathrm{B}-\mathrm{B}^{12}+\mathrm{B}^{13})y_t = (1+\alpha_1\mathrm{B}+\alpha_2\mathrm{B}^2+A_1\mathrm{B}^{12}+\alpha_1A_1\mathrm{B}^{13}+\alpha_2A_1\mathrm{B}^{14})\varepsilon_t$
=> $(1-\mathrm{B}-\mathrm{B}^{12}+\mathrm{B}^{13}-\rho_1\mathrm{B}+\rho_1\mathrm{B}^2+\rho_1\mathrm{B}^{13}-\rho_1\mathrm{B}^{14})y_t = (1+\alpha_1\mathrm{B}+\alpha_2\mathrm{B}^2+A_1\mathrm{B}^{12}+\alpha_1A_1\mathrm{B}^{13}+\alpha_2A_1\mathrm{B}^{14})\varepsilon_t$
=> $y_t - y_{t-1} - y_{t-12} + y_{t-13} - \rho_1y_{t-1} + \rho_1y_{t-2} + \rho_1y_{t-13} - \rho_1y_{t-14} = \varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + A_1\varepsilon_{t-12} + \alpha_1A_1\varepsilon_{t-13} + \alpha_2A_1\varepsilon_{t-14}$
=> $y_t = y_{t-1} + \rho_1y_{t-1} - \rho_1y_{t-2} + y_{t-12} - y_{t-13} - \rho_1y_{t-13} + \rho_1y_{t-14} + \varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + A_1\varepsilon_{t-12} + \alpha_1A_1\varepsilon_{t-13} + \alpha_2A_1\varepsilon_{t-14}$
=> $y_t = (1+\rho_1)y_{t-1} - \rho_1y_{t-2} + y_{t-12} - (1+\rho_1)y_{t-13} + \rho_1y_{t-14} + \varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + A_1\varepsilon_{t-12} + \alpha_1A_1\varepsilon_{t-13} + \alpha_2A_1\varepsilon_{t-14}$

(20 marks)

8. The SARIMA model `SARIMA.Ice.fit7` is a better model to use for prediction because although the RMSEP is larger (0.14 compared to the smallest of the other 3 models which was the Reduced Full Harmonic model 0.12) the acf of the SARIMA model does not have significant autocorrelation at lag 1 and the significant autocorrelation at lag 11 is small enough to ignore. So the assumptions of the SARIMA model are reasonably well satisfied and therefore the predictions should be reliable.

(10 marks)

9. sim1: ARMA(1,1) as we have decay in the acf and in the pacf but no indication of the order required so we begin with ARMA(1,1)
$y_t = \rho_1y_{t-1} + \varepsilon_t + \alpha_1\varepsilon_{t-1}$

sim2: AR(2) as we have decay in the acf and cut-off at lag 2 in the pacf
$y_t = \rho_1y_{t-1} + \rho_2y_{t-2} + \varepsilon_t$

sim3: MA(2) as we have cut-off at lag 2 in the acf and persistence in the pacf
$y_t = \varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2}$

sim4: White Noise as there are no significant lags in the acf or the pacf so the series is iid N(0, σ^2)
$y_t = \varepsilon_t$

(20 marks)