Zoe Zhou 269999947

326 Test 2

1. The plot of the Global Nitrous Oxide data showed a strong positive linear increasing trend, with reasonably constant seasonal components pattern. But in year 2014, the seasonal component flattened down slightly. The linear trend seems to be very steady and the global N2O has increased from 316 to 332 ppb, with an increase of approximately 1 ppb per year.

   There might be a change of slope (a break) somewhere in the series that caused the inconsistent seasonality. The seasonal peak seemed to be in January and the lowest showed in August. It is very likely to have autocorrelation in this series.

2. The residual series seemed to follow a normal distribution with mean 0 and constant variance. The p-value from Shapiro-walk test is 0.99 meaning there is no evidence against the null hypothesis that our residuals are from a normal distribution. Checking the 5-value of the residuals in the summary output, we can say the residuals is normal.

   However, the plot of ACF showed we have lag (1) being significant. It is relatively large we cannot ignore it. So, we still have autocorrelation problem. We do not meet the assumptions since we still have autocorrelation issues.

3. Prediction for 2019.7:

   31.7838298 + 0.0061339 * 223 + 0.0018856 * (223 - 80) - 0.1121256 + 0.8996494 * 331.7

   = 331.7229107 = 331.72 ppb

   Prediction for 2019.8:

   31.7838298 + 0.0061339 * 224 + 0.0018856 * (224 - 80) - 0.0677393 + 0.8996494 * 331.7229107

   = 331.7959281 = 331.80 ppb

   RMSEP:

   sqrt (1/4 * [$(331.9 - 331.7229107)^2$ + $(331.9 - 331.7959281)^2$ + $(331.9 – 331.9300)^2$ + $(331.9 - 332.0974)^2$])

   = 0.098994 ppb

4. If we fit a Full Harmonic model, the predictions and RMSEP will be the exact same as the Seasonal Factor model, so same as results in the last question.

5. The predictions from two of the Reduced Harmonic models are really close to our prediction from the Seasonal Factor model, with a difference of around 0.01 ppb. They all very similar to the actual values.

   The RMSEP are all very similar. But the RMSEP from the Significant Harmonics model is the smallest (0.08951), and the other two models are slightly higher by about 0.01 ppb (0.09899 and 0.09485).

6. The Full Harmonic model with significant pairs has harmonics c1, s1, c2 and s2. The Full Harmonic model with only significant harmonics only has harmonics c1, s1 and c2.

7. Since we suspect the seasonal peak occurs in January. The model formula is:

   seasonal = cos ((2 pi * (red.Time - 1)) / 12)

   cosine.fit = red.N2O.ts[-1] ~ red.Time[-1] + red.Time.break[-1] + seasonal[-1] + red.N2O.ts[-222])

   The prediction from the Cosine model is very close to the previous models we have tried (The Full and Reduced Harmonic Models). They are very similar to the actual values. But the RMSEP of the Cosine model is the largest among all (0.0966). So, the Cosine model is not very useful here, we wouldn't prefer Cosine model.

8. TS1: AR(2)

   Because we have decay in ACF and a cut-off at 2 significant lags in PACF.

   Model equation: $y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t \quad where\ \varepsilon_t \sim iid\ N\,(0,\ \sigma^2)$

   Other models to try: AR(3), ARMA(1, 1), ARMA(1, 2), ARMA(2, 1)

   TS2: MA(2)

   Because we have cut-off in ACF at lag(2) and some persistence pattern in PACF.

   Model equation: $y_t = \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} \quad where\ \varepsilon_t \sim iid\ N\,(0,\ \sigma^2)$

   Other models to try: MA(3), ARMA(1, 1), ARMA(1, 2), ARMA(2, 1)

TS3: WN

Because we see no significant lags in ACF or PACF. The series showed very normal pattern with constant variance so this is a white noise series.

Model equation: $y_t = \varepsilon_t \quad where\ \varepsilon_t \sim iid\ N\ (0,\ \sigma^2)$

Other models to try: None

TS4: ARMA(1, 1)

Because we have decay in both ACF and PACF. We don't know the order of the model so we should start trying from ARMA(1, 1).

Model equation: $y_t = \rho_1\, y_{t-1} + \varepsilon_t + \alpha_1\, \varepsilon_{t-1} \quad where\ \varepsilon_t \sim iid\ N\ (0,\ \sigma^2)$

Other models to try: ARMA(1, 2), ARMA(2, 1), ARMA(2, 2)…