SARIMA fitting: Milk production

PRACTICAL TIME SERIES ANALYSIS

THISTLETON AND SADIGOV

Objectives

- Fit SARIMA models to Milk production data from TSDL
- Forecast future values of examined time series

Modeling

- Time plot
- Transformation
- Differencing (seasonal or non-seasonal)
- ACF → Adjacent spikes → MA order
- ACF → Spikes around seasonal lags → SMA order
- PACF → Adjacent spikes → AR order
- PACF → Spikes around seasonal lags → SAR order

Modeling cont.

- Fit few different models
- Compare AIC, choose a model with minimum AIC
- **The parsimony principle**
- Time plot, ACF and PACF of residuals
- Ljung-Box test for residuals

The parsimony principle

SARIMA$(p, d, q, P, D, Q)_s$

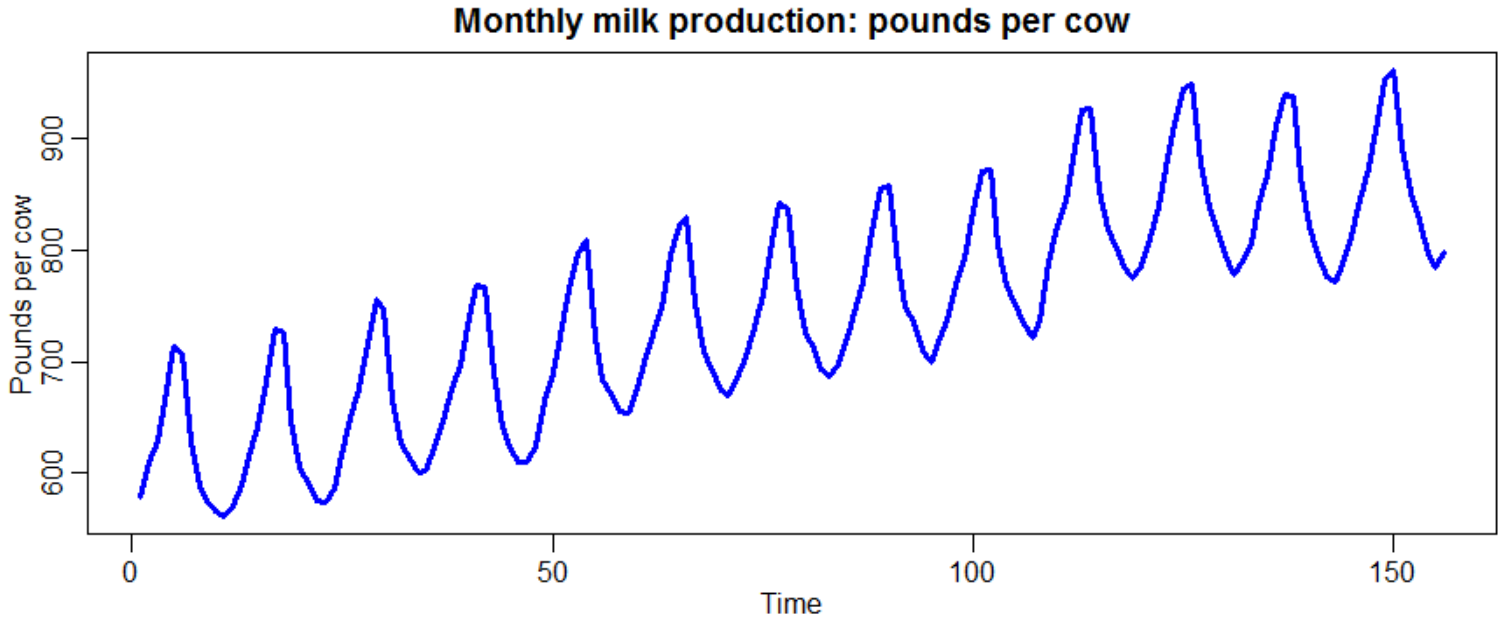
$$p + d + q + P + D + Q \leq 6$$

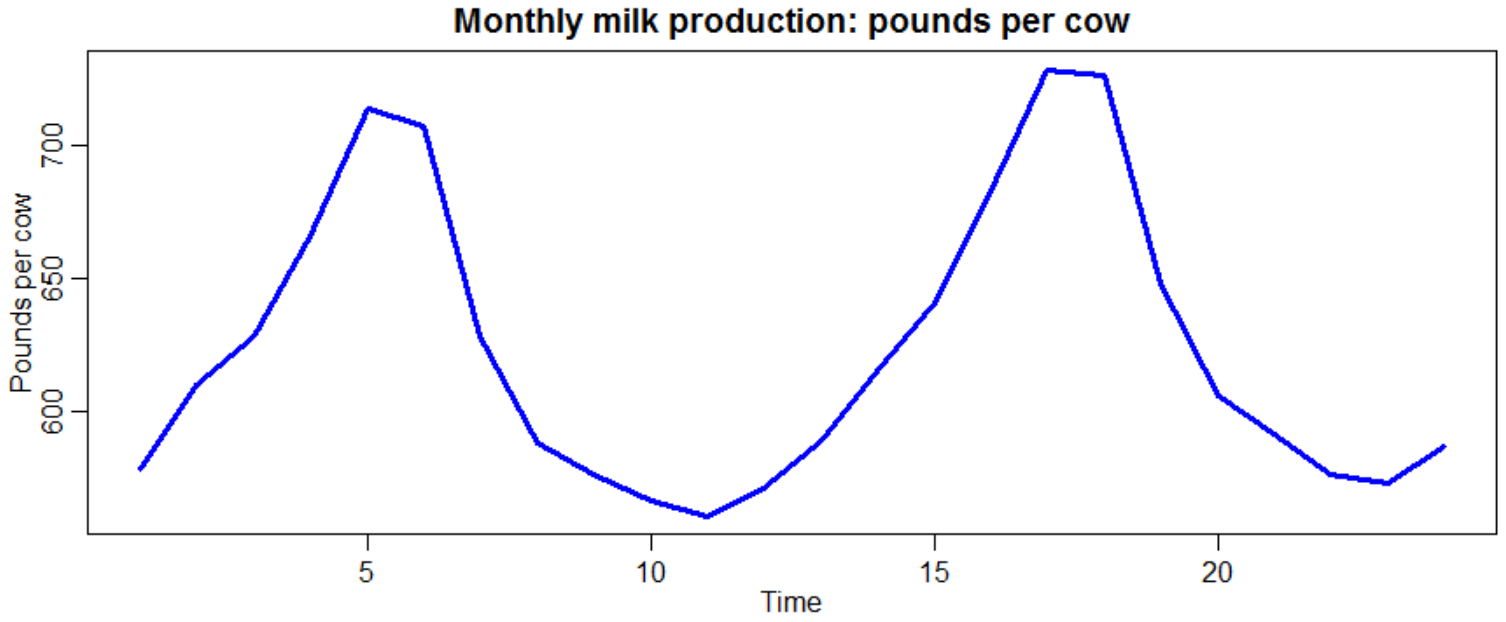
Time Series Data Library

- TSDL
- Created by Rob Hyndman
- Professor of Statistics
- Monash University, Australia
- https://datamarket.com/data/list/?q=provider%3Atsdl

Monthly milk production : Agriculture

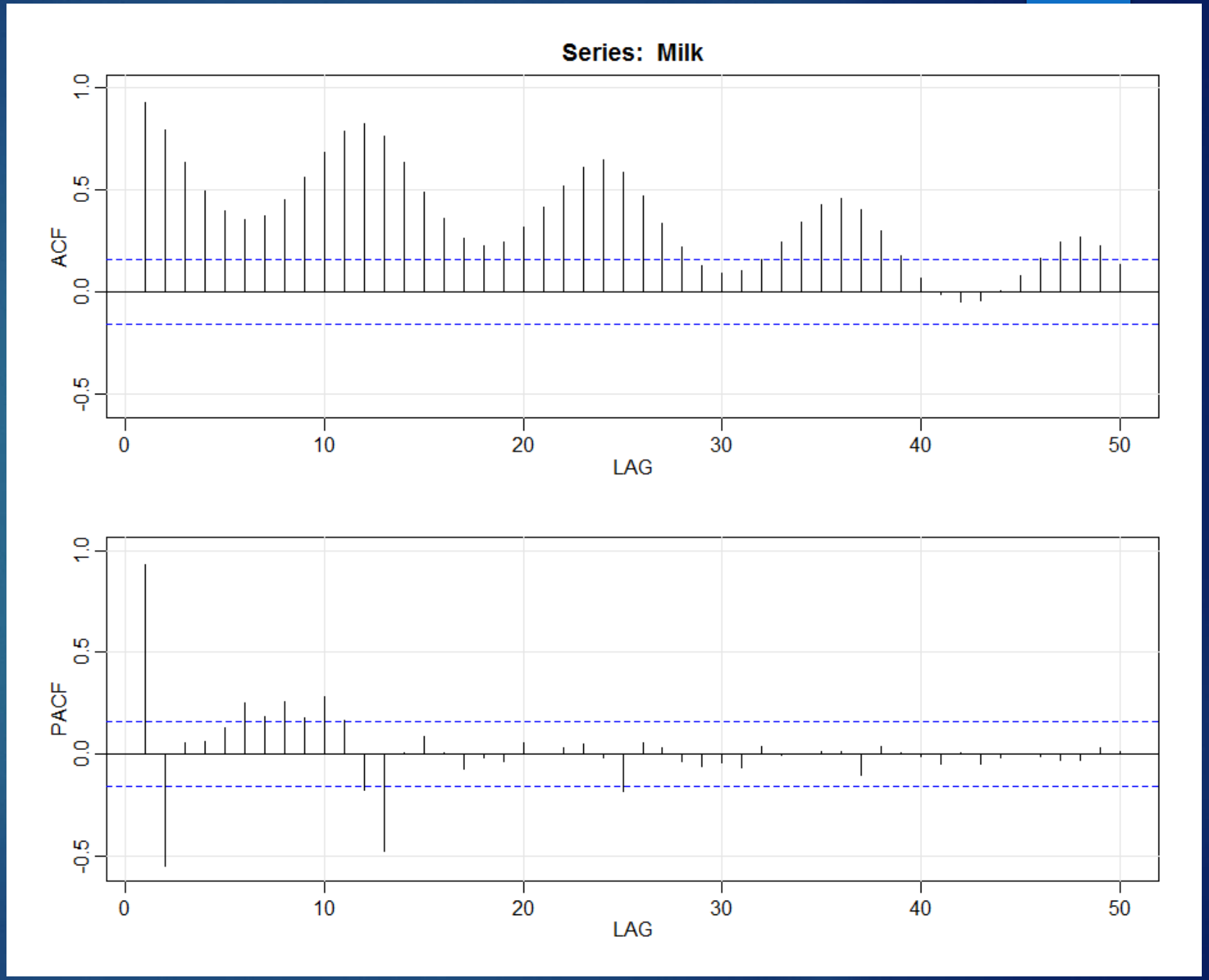
- https://datamarket.com/data/set/22sn/monthly-milk-production-pounds-per-cow-jan-62-dec-75-adjusted-for-month-length#!ds=22sn&display=line
- Monthly milk production
- Pounds per cow
- January 1962 – December 1975
- Agriculture, Source: Cryer (1986)


Monthly milk production: pounds per cow
Pounds per cow
600
700
800
900
0
50
100
150
Time


Monthly milk production: pounds per cow
Pounds per cow
600
650
700
5
10
15
20
Time

ACF

PACF

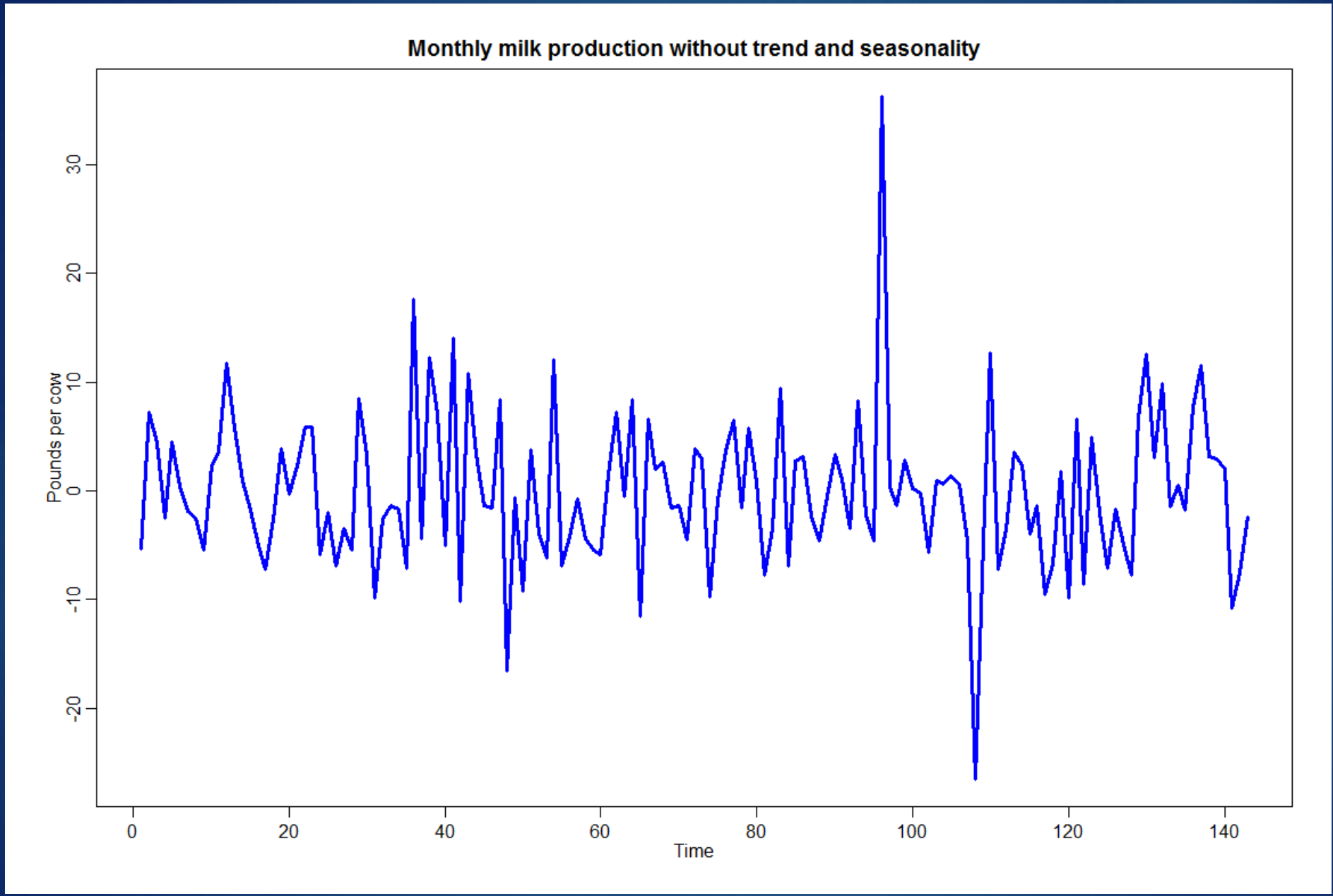

Series: Milk
ACF
1.0
0.5
0.0
-0.5
0
10
20
30
40
50
LAG
PACF
1.0
0.5
0.0
-0.5
0
10
20
30
40
50
LAG

Non-seasonal and seasonal differencing

$$d = 1$$

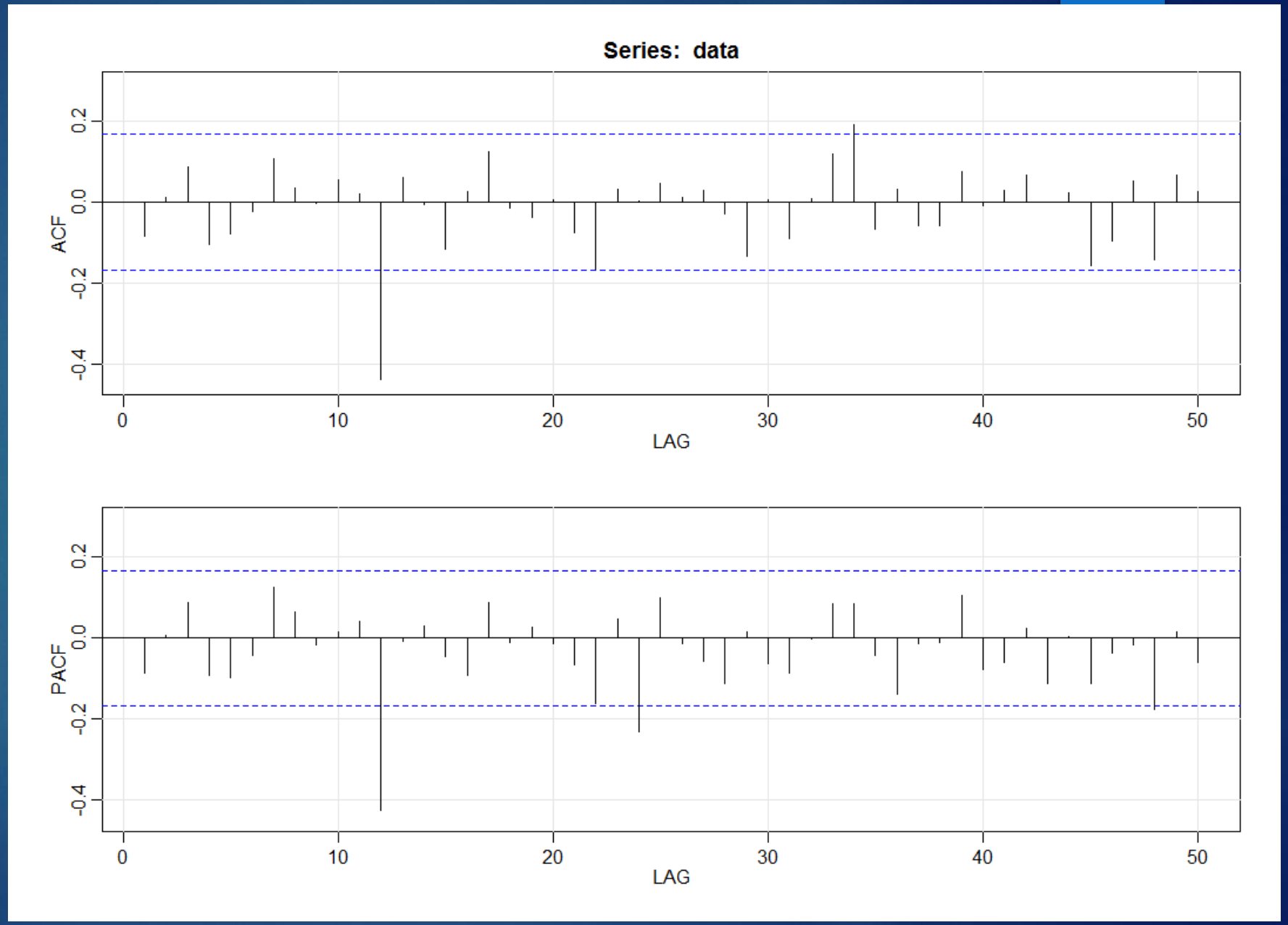
$$D = 1$$

$$diff(diff(milk), 12)$$


Monthly milk production without trend and seasonality
Pounds per cow
30
20
10
0
-10
-20
0
20
40
60
80
100
120
140
Time

ACF

PACF


Series: data
ACF
0.2
0.0
-0.2
-0.4
0
10
20
30
40
50
LAG
PACF
0.2
0.0
-0.2
-0.4
0
10
20
30
40
50
LAG

Order specification

- ACF → $q = 0$; $Q = 0, 1, 2, 3$
- PACF → $p = 0$; $\mathrm{P} = 0$, $1, 2$

0 1 0 0 1 0 12 AIC= 968.3966 SSE= 7213.013 p-VALUE= 0.4393367

0 1 0 0 1 1 12 AIC= 923.3288 SSE= 4933.349 p-VALUE= 0.6493728

0 1 0 0 1 2 12 AIC= 925.3072 SSE= 4931.398 p-VALUE= 0.6529998

0 1 0 0 1 3 12 AIC= 927.2329 SSE= 4925.911 p-VALUE= 0.6640233

0 1 0 1 1 0 12 AIC= 938.6402 SSE= 5668.197 p-VALUE= 0.493531

0 1 0 1 1 1 12 AIC= 925.3063 SSE= 4931.428 p-VALUE= 0.6531856

0 1 0 1 1 2 12 AIC= 927.3036 SSE= 4931.135 p-VALUE= 0.6537708

0 1 0 1 1 3 12 AIC= 929.2146 SSE= 4924.747 p-VALUE= 0.6627108

0 1 0 2 1 0 12 AIC= 932.6438 SSE= 5308.012 p-VALUE= 0.6004804

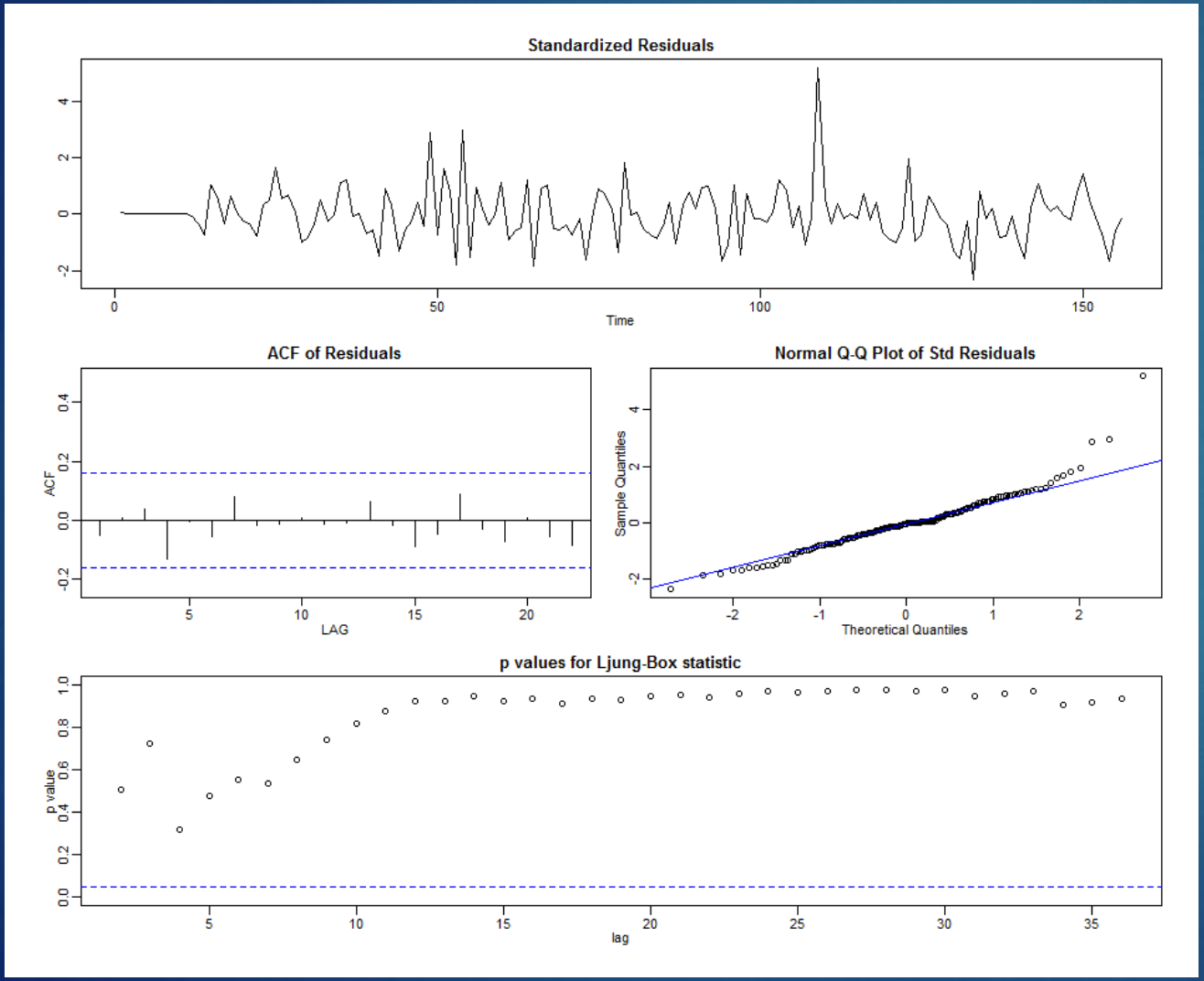
0 1 0 2 1 1 12 AIC= 927.2797 SSE= 4929.733 p-VALUE= 0.657349

0 1 0 2 1 2 12 AIC= 926.8053 **SSE= 4618.498** p-VALUE= 0.6826743

$SARIMA(0,1,0,0,1,1)_{12}$

	Estimate	SE	t.value	p.value
sma1	-0.6750	0.0752	-8.9785	0.0000

Residual analysis



Model – $SARIMA(0,1,0,0,1,1)_{12}$

$$X_t = Milk\ production\ pounds\ per\ cow$$

$$(1 - B)(1 - B^{12})X_t = (1 + \Theta B^{12})Z_t$$

$$X_t = X_{t-1} + X_{t-12} - X_{t-13} + Z_t + \Theta\, Z_{t-12}$$

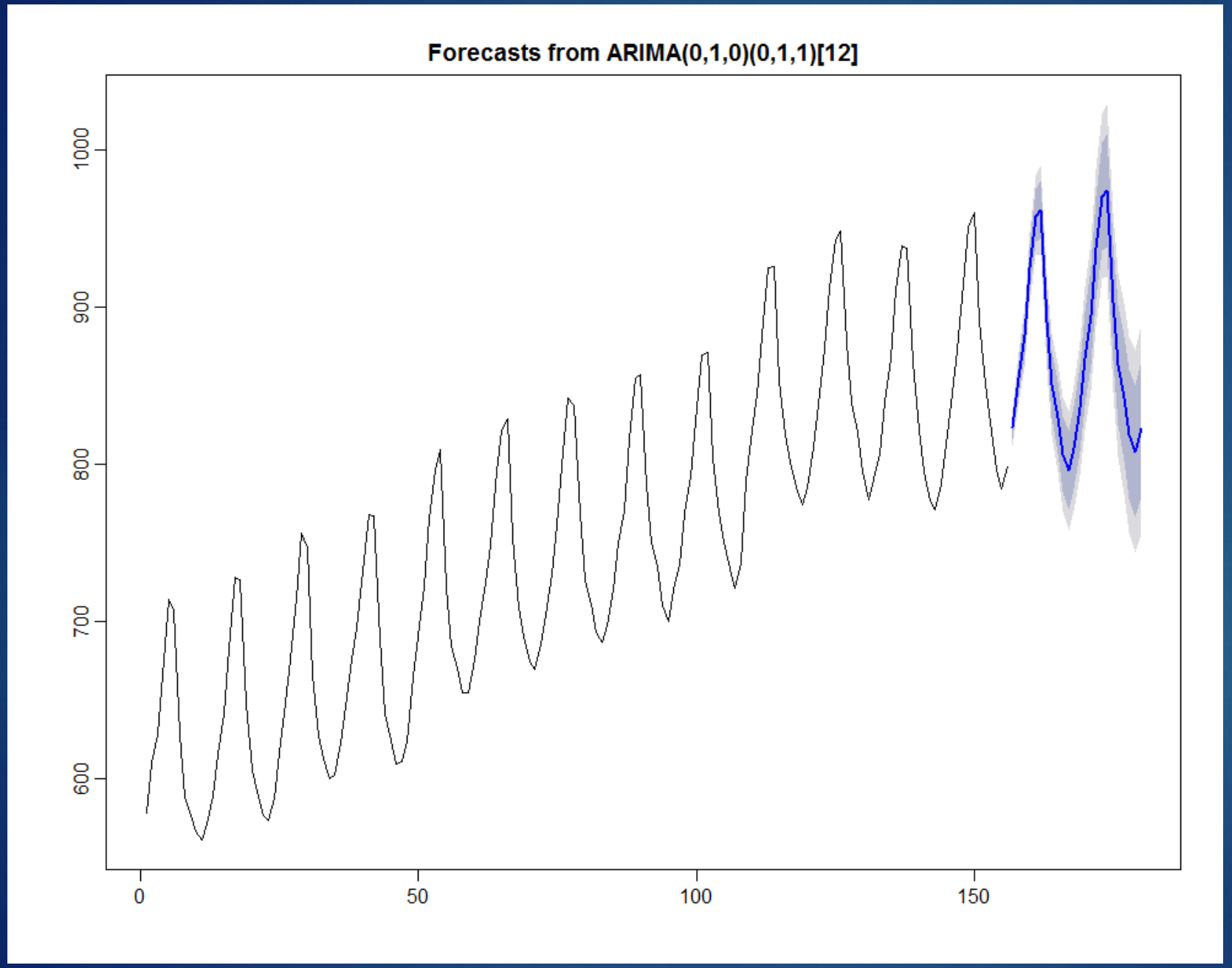
$$\widehat{\Theta} = -0.6750$$

Model – cont.

$$X_t = X_{t-1} + X_{t-12} - X_{t-13} + Z_t - 0.6750\, Z_{t-12}$$

where

$$Z_t \sim Normal\ (0, 34.47)$$


Forecasts from ARIMA(0,1,0)(0,1,1)[12]
1000
900
800
700
600
0
50
100
150

forecast(model)

	Pt. for.	Lo 80	Hi 80	Lo 95	Hi 95
157	823.3978	815.8740	830.9216	811.8911	834.9045
158	854.9196	844.2793	865.5598	838.6467	871.1925
159	882.1923	869.1607	895.2239	862.2622	902.1224
160	925.2390	910.1914	940.2866	902.2257	948.2523
161	958.4461	941.6225	975.2698	932.7165	984.1757
162	962.2105	943.7811	980.6399	934.0252	990.3959
163	890.9973	871.0912	910.9033	860.5536	921.4409
164	851.3336	830.0531	872.6140	818.7879	883.8792
165	829.7513	807.1800	852.3226	795.2314	864.2711
166	806.7802	782.9880	830.5725	770.3931	843.1673
167	795.9513	770.9978	820.9048	757.7882	834.1144
168	810.5435	784.4804	836.6066	770.6834	850.4036

What We've Learned

- Fit SARIMA models to Milk production data from TSDL
- Forecast future values of examined time series