Forecasting Milk Production

Nádia Soares

Séries Temporais e Previsão

Faculdade de Ciências da Universidade do Porto

January 14th, 2018

Summary

Introduction

- Agriculture and livestock sectors are vital industries for several communities
- For some, milk production accounts for more than 20% of the total agricultural value
- It provides a significant amount of exportation and jobs to those communities
- Forecasting its production can help planning and managing operations accordingly

Objectives

- Exploratory and describe a milk production dataset
- Model the time series and provide forecasts
- Experiment and compare different methods, namely, ARIMA and exponential smoothing
- Present the results

The time series

- Milk production from Canada, in pounds per cow
- 19 years of monthly data (228 observations): from January 1995 to December 2013
- ⟨https://raw.githubusercontent.com/ricardoscr/UW-Data-Science-Certificate/master/02-Methods/CADairyProduction.csv⟩

The time series

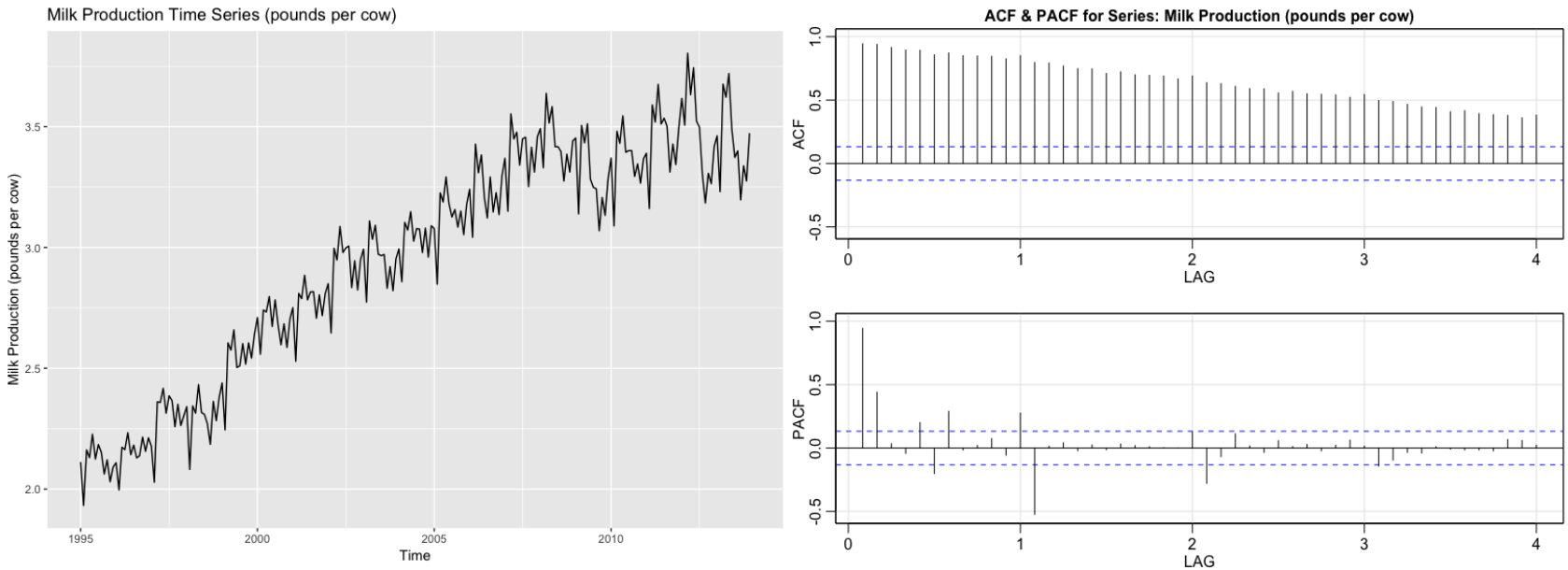


Figura 1: a) Values of milk in pounds per cow plotted over time. b) Autocorrelation and partial autocorrelation functions.

The time series

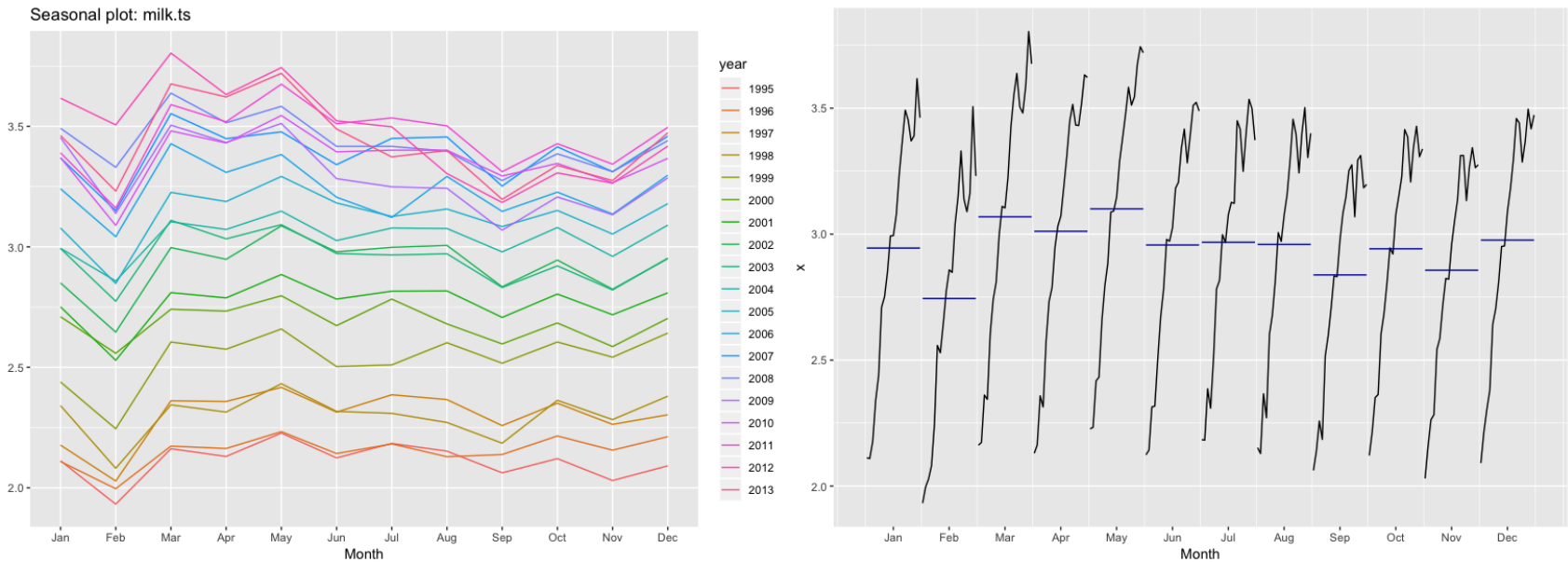


Figura 2: a) Season plots and b) month plot. Both plots illustrate an increase trend and a season effect. Values decrease in February, September and November and pick in March and May.

Removing Heteroscedasticity

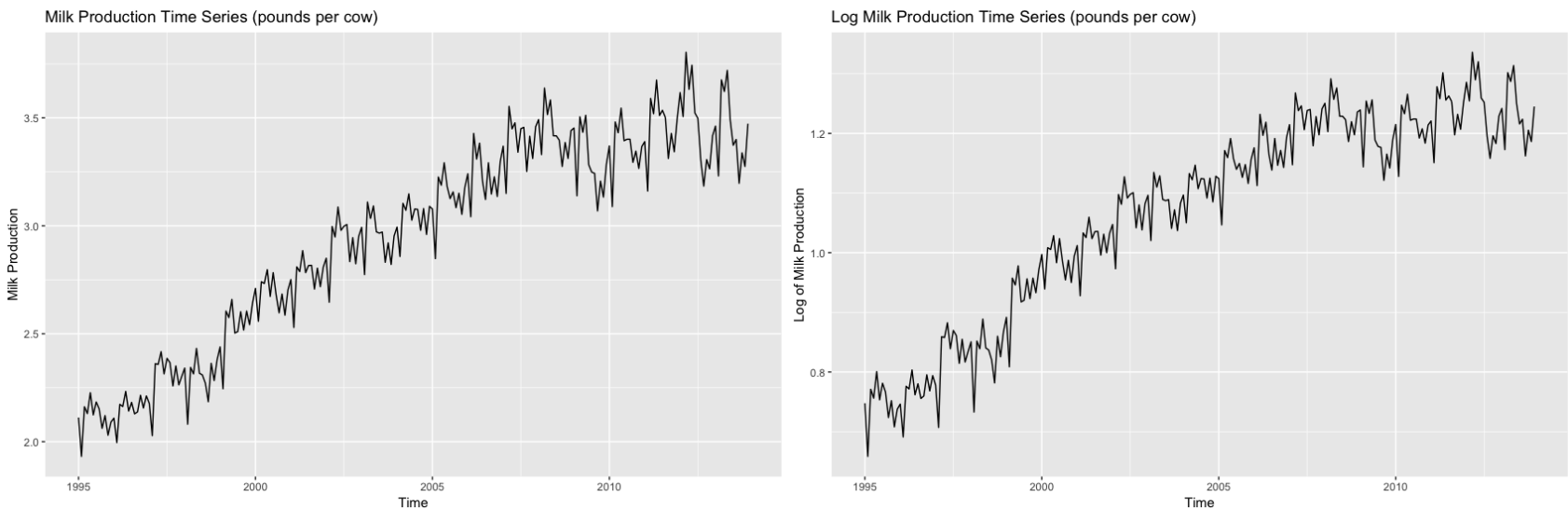


Figura 3: a) Original and b) logged milk production time series

Removing Trend - 1st difference

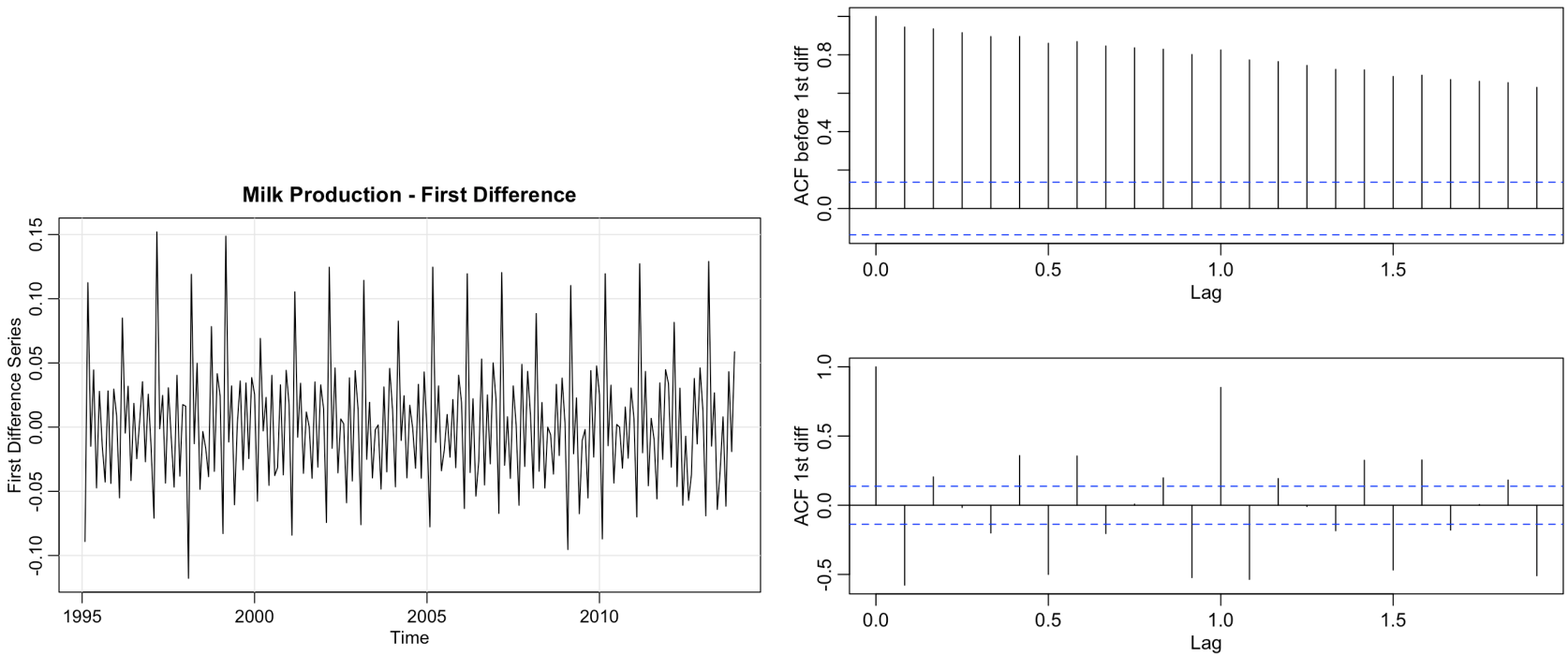


Figura 4: a) First difference of the logged time series and b) its ACF and PACF

Removing Seasonality - 1st and 12th difference

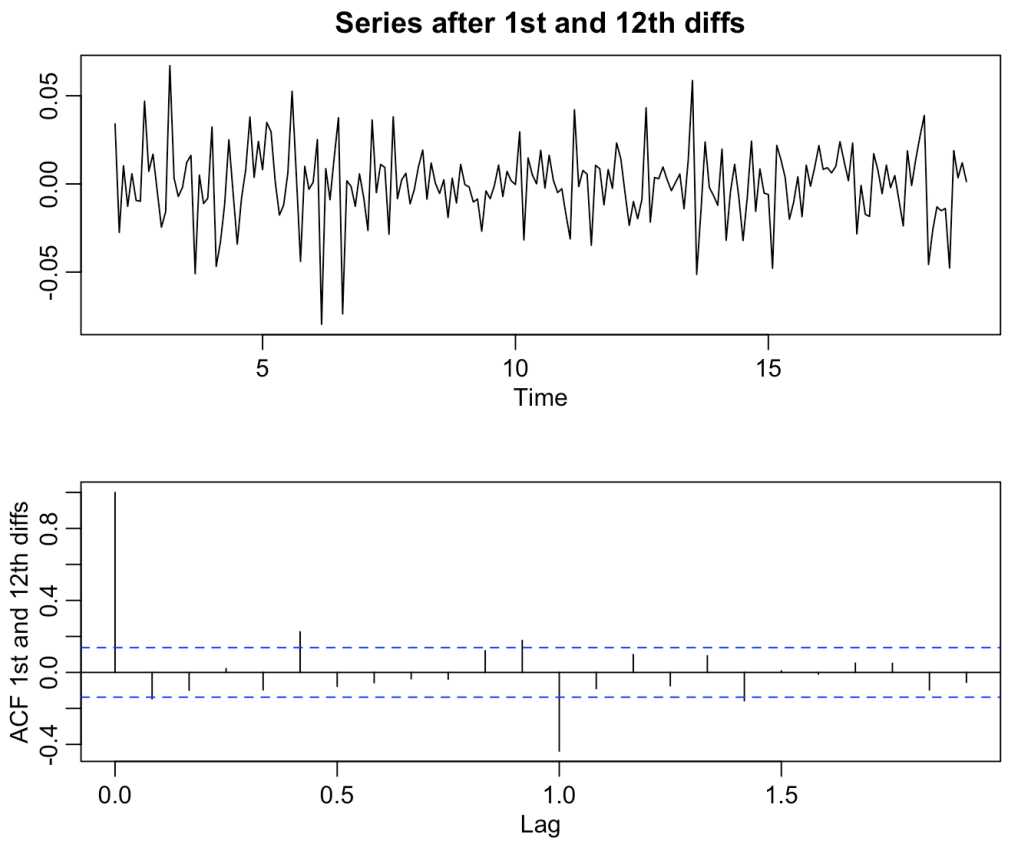


Figura 5: a) Both 1st and 12th differences and b) the ACF of the result

ACF and PACF

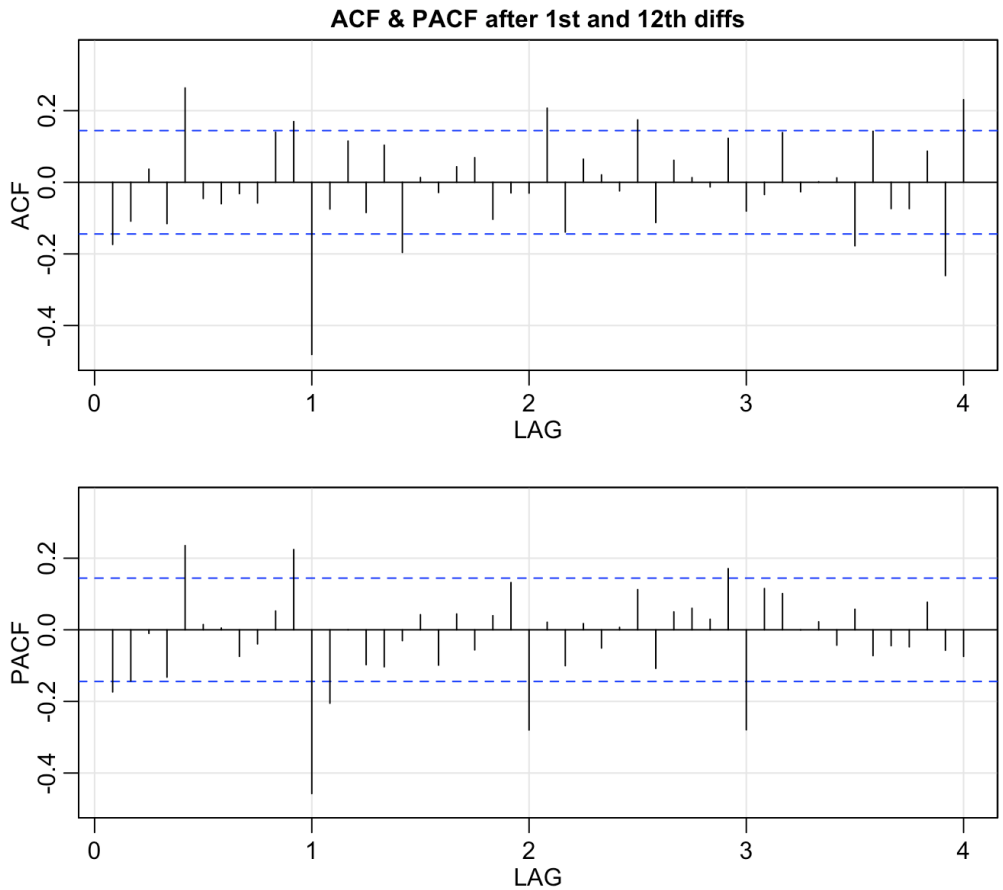


Figura 6: ACF and PACF of the differenciated series

SARIMA

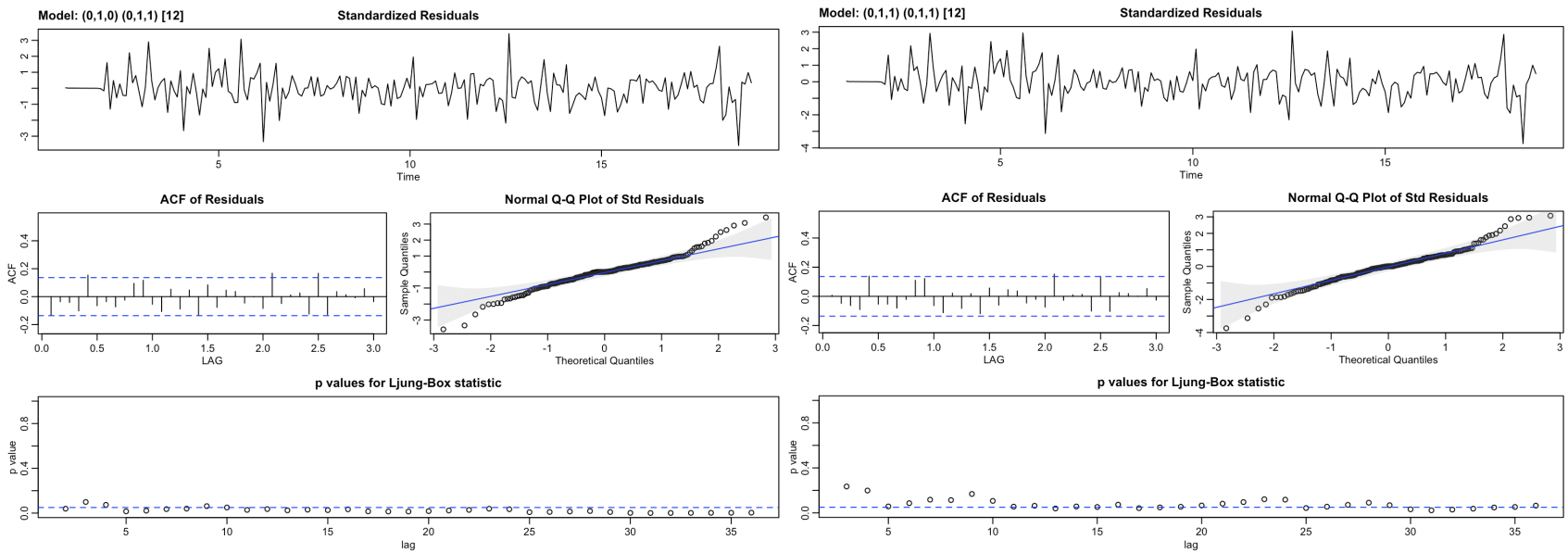


Figura 7: Result applying a a) SARIMA(0,1,0)(0,1,1)[12] and a b) SARIMA(0,1,1)(0,1,1)[12]

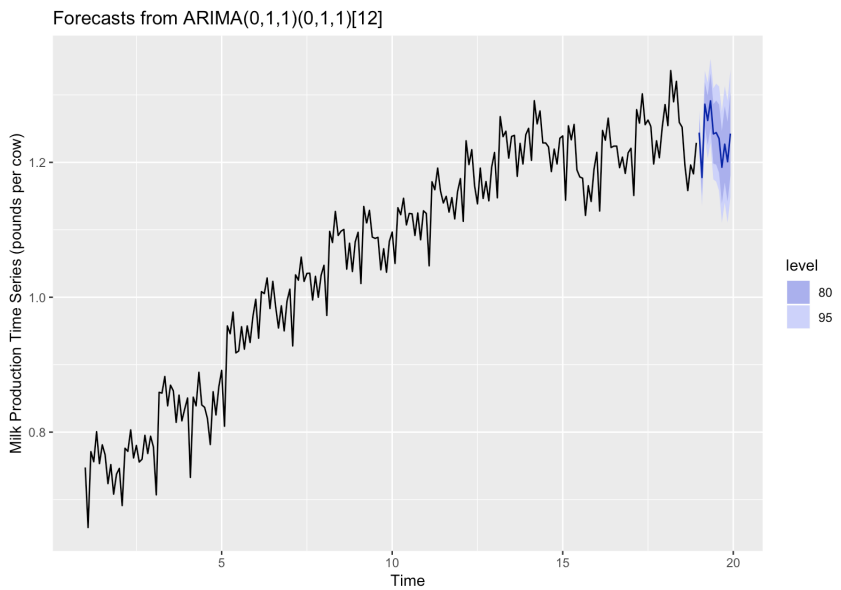
SARIMA

(p,d,q)(P,D,Q)	AIC	AICc	BIC	non-significant coefficients
(0,1,0)(0,1,1)	-7.247	-7.238	-8.232	no
(0,1,1)(0,1,1)	-7.253	-7.243	-8.221	no
(0,1,2)(0,1,1)	-7.246	-7.236	-8.200	yes
(0,1,5)(0,1,1)	-7.242	-7.230	-8.148	yes
(1,1,0)(0,1,1)	-7.251	-7.242	-8.220	yes
(5,1,0)(0,1,1)	-7.246	-7.235	-8.153	yes
(1,1,1)(0,1,1)	-7.252	-7.242	-8.205	yes

Tabela 1: Criteria of different models tested

SARIMA

The time series was divided into train and test set. The test set is the last 1 year.



	RMSE	MAE	MPE	MAPE
SARIMA(0,1,1)(0,1,1)[12]	0.0183	0.0155	-0.3215	1.2624

Exponential Smoothing - Holt-Winters

	AIC	AICc	BIC
Additive	-558.17	-555.07	-500.79
Multiplicative	-517.18	-514.09	-459.80

Tabela 2: AIC, AICc and BIC values for additive and multiplicative Holt-Winters

Exponential Smoothing - Holt-Winters

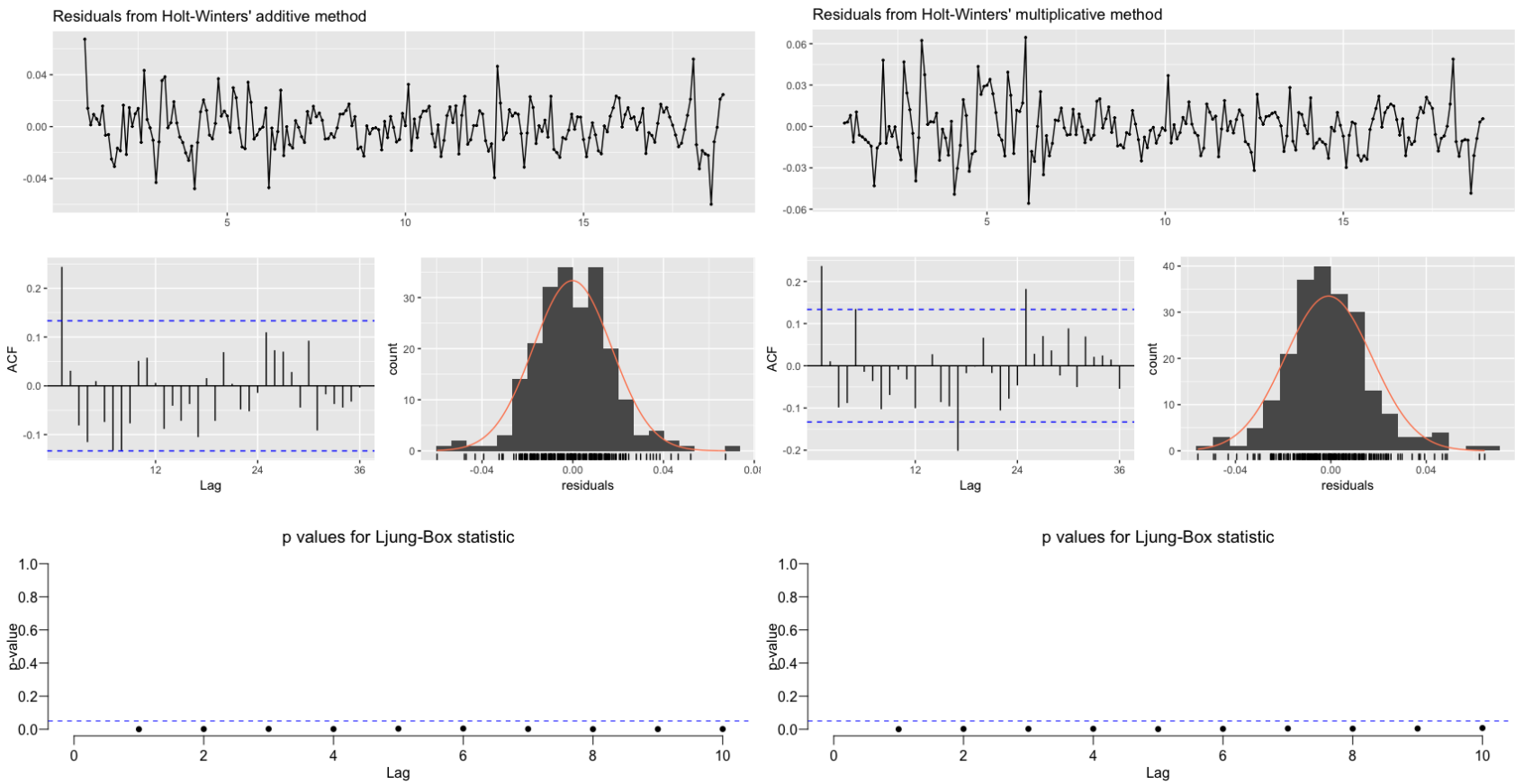
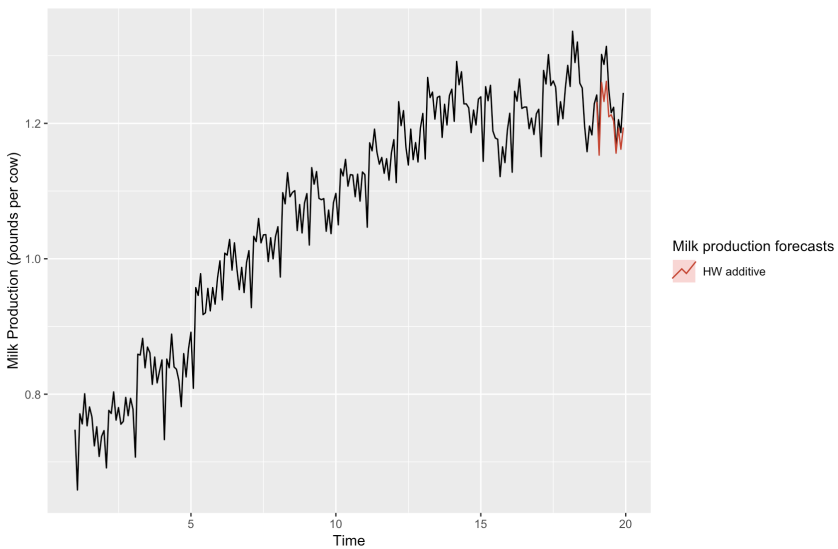
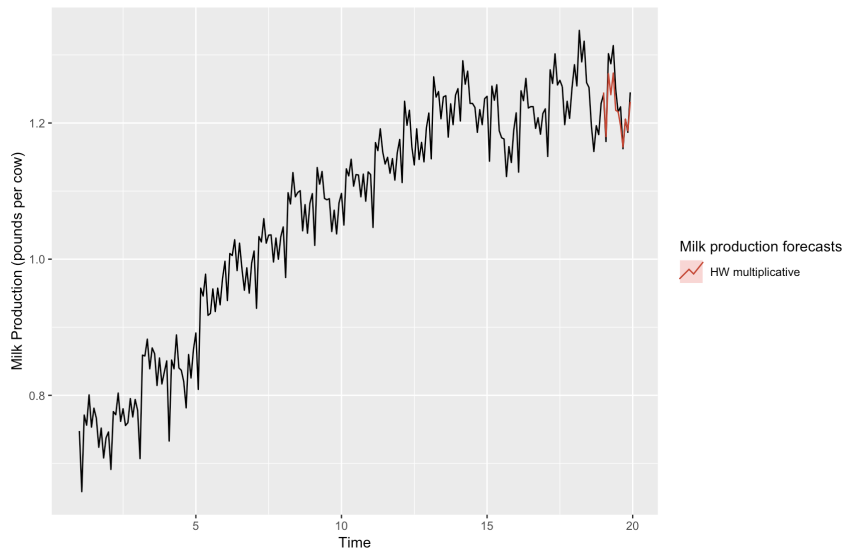


Figura 8: Residuals of additive and multiplicative Holt-Winters

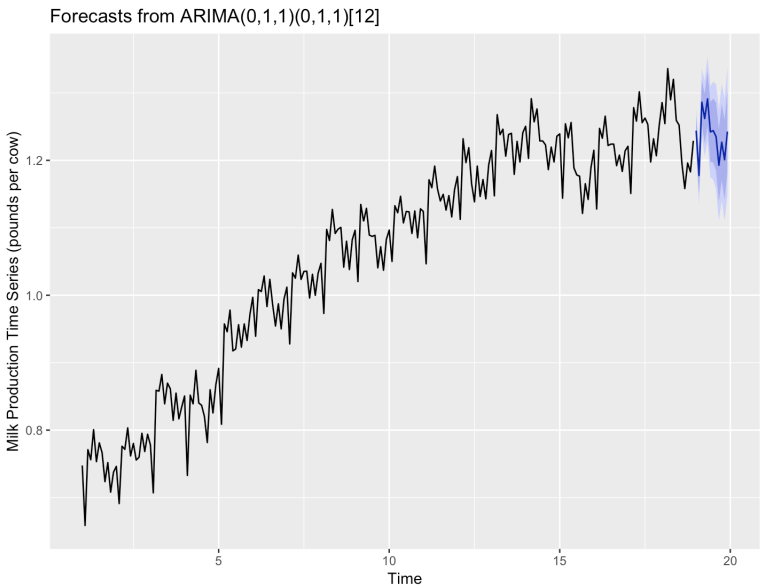
Exponential Smoothing - Holt-Winters

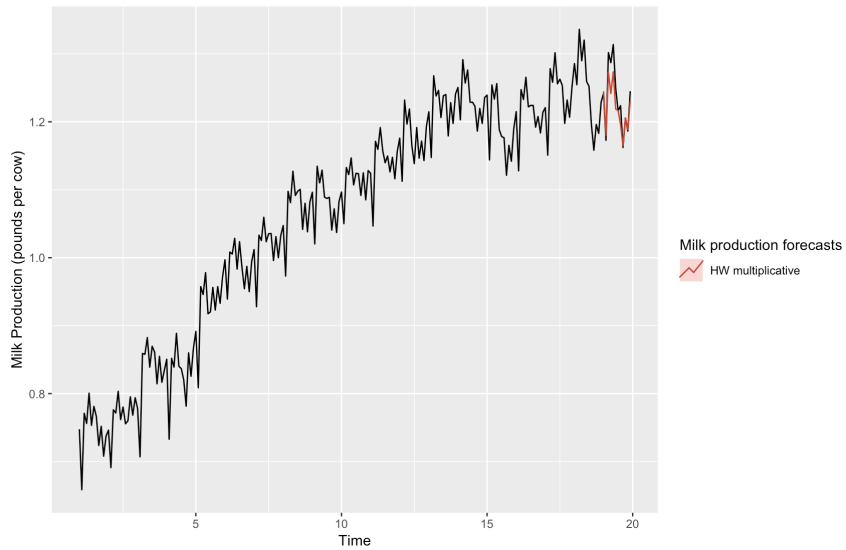




	RMSE	MAE	MPE	MAPE
Additive	0.0333	0.0302	-1.4265	2.4353
Multiplicative	0.0286	0.0252	-1.6268	2.0318

Results - SARIMA vs Holt-Winters





	RMSE	MAE	MPE	MAPE
SARIMA(0,1,1)(0,1,1)[12]	0.0183	0.0155	-0.3215	1.2624
Multiplicative HW	0.0286	0.0252	-1.6268	2.0318

Conclusion

- In this particular dataset, a SARIMA model had the best results
- SARIMA models also creates intervals while Holt-Winters only generates point forecasts
- In the future it would be interesting to do a multivariant analysis using the milk and the ice cream data present in the original dataset

References

1 IDF. IDF Factsheet-February 2013 Scientific excellence Industry applicability Strategic net-working Global influence Fig. 2: Share of milk in total agricultural production-situation in 2010according to countries. (February), 2013.

2 OECD and FAO. Dairy and dairy products.Agricultural outlook 2018-2027, page 12, 2018.

3 B Blaskó. World Importance and Present Tendencies of Dairy Sector.Production, 2010.

4 https://otexts.org/fpp2/