Université catholique de Louvain
Louvain School of Statistics

LSTAT2170 - Time series

Final Project

Lionel Lamy
Noma: 1294-1700

April 8, 2021

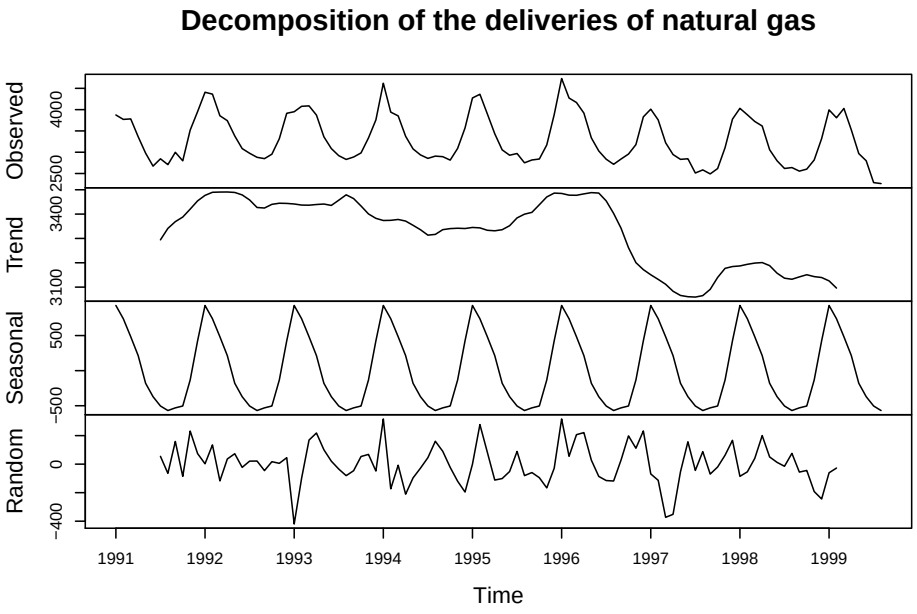
Contents

1 Introduction

In this project we will focus on the analysis of real data. These data are provided by the U.S. Natural Gas State Data and concern the aggregated monthly quantity of natural gas delivered to residential and commercial consumers (excluding vehicle fuel) in Florida. Aggregated on a monthly basis, the data are presented in millions of cubic feet (MMcf[1]) and cover the period from January 1991 to August 1999.

We will proceed to the beginning of this report with a first visual discovery of the dataset. Then, we will eventually apply a series of transformations to stabilize the variance, remove possible trends and seasonality. Next, we will analyze the autocorrelation and partial autocorrelation functions in order to have a first intuition on the type of model to fit. Following this, we will establish which model would be the most appropriate for our data and verify the insight of our choice by several methods such as a significance test of the coefficients, an analysis of the residuals (by a Portmanteau test) or the evaluation of the predictive capacity "on sample". The final objective is to be able to give a prediction interval for future values over roughly one year.

2 Data discovery

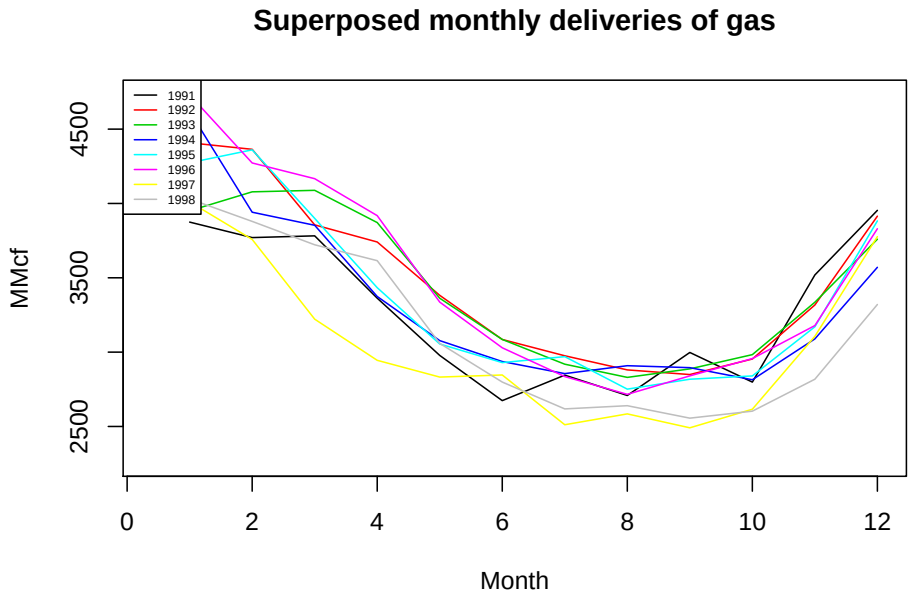


In the graph above, we can observe an almost perfect seasonality with maximum values during the months of January and minimum from June to August. These results are hardly surprising given the nature of the data. Indeed, it seems normal that a greater quantity of gas is used in January (as it it the coldest month in Florida[2]) and that gas consumption decreases during summer.

Looking at the trend line, we notice that from 1996 onwards the quantity delivered seems to decrease. We will try to keep this in mind when trying to predict future values.

[1] "Mcf" means 1,000 cubic feet of natural gas; "MMcf" means 1,000 Mcf.

[2] Source: weather-us.com[3]



3 Model selection

3.1

3.2

4 Conclusion

A Appendix

A.1 Code

Note

For reproducibility purposes, the complete R project containing the source code and the results is available on https://github.com/lamylio/LSTAT2170-Project.

```r
# Import facility functions
source("./resources/scripts/fonctionsSeriesChrono.R")

source("./resources/scripts/custom_plots.R")
source("./resources/scripts/custom_timeseries.R")

# Import the gas dataset
gas = read.table("./resources/data/gasflorida.txt", header = F)
gas = ts(gas, start=1991, frequency = 12)

# Define some useful variables
gas.attributes = attributes(gas)$tsp
gas.start = gas.attributes[1]
gas.end = gas.attributes[2]
gas.freq = gas.attributes[3]
```