EcoDep course 2023

Exam

Introduction to univariate time series analysis for climate data

Teacher: Federico Maddanu *

Instructions: Please, answer to the following questions in a pdf file. Then, send it along to a .zip folder with the relative OCTAVE code to the mail address federico.maddanu@gmail.com using as object of the mail EXAM ECODEP 2023 and providing the name, surname and ID of the candidate in the main body. Mails must be send before the end of March 2023.

1. Consider the MA(2) process $y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2}$ with $\varepsilon_t \sim WN(0, \sigma^2)$.

   i) Compute the ACF and show that $\gamma(k) = 0$ for $k > 2$. Show that $f(\omega) = \frac{\sigma^2}{2\pi}[1 + \psi_1^2 + \psi_2^2 + 2(\psi_1 + \psi_1\psi_2)\cos(\omega) + 2\psi_2 \cos(2\omega)]$.

   Hints: you can use the following relations: $\imath = \sqrt{-1}$, $e^{-\imath\omega} = \cos(\omega) - \imath \sin(\omega)$, $|a + \imath b| = \sqrt{a^2 + b^2}$ with $a, b \in \mathbb{R}$ and trigonometric identities.

   ii) Simulate the process in OCTAVE considering $\sigma^2 = 1$, $\psi_1 = 0.5$, $\psi_2 = -1.2$ and $n = 1000$. Plot the series, the sample autocorrelation, the sample partial correlation and in the same graph plot the periodogram along to the theoretical spectral density. Why we should deduce from the correlogram that the we have a MA(2) process?

2. Implement a Monte Carlo experiment with 1000 reps. For each reps:

   i) First simulate $n = 400$ realizations from the random process with occasional shifts in the mean:

   $$y_t = \varepsilon_t + \sum_{i=1}^{t} \eta_i b_i \tag{1}$$

   where $\varepsilon_t \sim N(0, 1)$, $\eta_t \sim N(0, 0.1)$, $E(\varepsilon_t \eta_t) = 0$ and b_i follows an i.i.d. binomial distribution with $p = \text{Prob}(b_i = 1)$ representing the probability of a break and η_t establishes the size of its jump, such that pn returns the expected number of breaks. (**Hints**: use the function 'binornd()' after loading it by running the code 'pkg load statistics').

   ii) Then estimate at each reps the FN model and the memory parameters d.

*Postdoc researcher at the EcoDep project, AGM Laboratory, CY University, 2 av. Adolphe Chauvin (Bat. E, 5th floor), CERGY-PONTOISE, 95302, France, mail: federico.maddanu@gmail.com

Finally, compute the Monte Carlo mean $E(\hat{d}) = \frac{1}{10^3}\sum_{j=1}^{10^3} \hat{d}_j$ of the memory parameters on all the 1000 reps. Repeat the experiment for the values of $p \in \{0.05, 0.15, 0.25\}$.

What are the conclusions of the experiments as pn increases?

3. Figure 1 shows the weekly mean of the ground-based Fourier Transform Infrared (FTIR) solar spectra measurements of ethane (C_2H_6) recorded at the ground-station of Toronto (Canada). Ethane is the most abundant non-methane hydrocarbon in the atmosphere. Understanding its dynamics is of crucial importance in the context of climate change. Indeed, ethane affects the distribution of ozone (O_3) and even though the ozone layer protect the life in the planet absorbing most of the Sun's ultraviolet radiation, the formation of ground level ozone has pollution effects on the air quality and damages ecosystems. Furthermore, ethane is an indirect greenhouse gas, which influences the atmospheric lifetime of methane (CH_4), such that ethane emissions can be used as a measure of methane emissions. The economic exploitation of oil and natural gas is the main driver of the co-emission of the two gases. However, while methane is released in the atmosphere by both natural events (as biomass burning emissions) and anthropogenic activities, ethane emissions do not have significant natural sources. It has been reported in the literature (see Franco et al., 2015, Franco et al., 2016, Helmig et al., 2016 and the references therein) that the levels of ethane emission can be most likely attributed to the production and transport of oil and natural gas in United States (US).

   i) Upload the data 'TorontoC2H6.csv' in OCTAVE by running:

```
Data=importdata("TorontoC2H6.csv");
y=Data.data;
dates=2000 + datenum(Data.textdata(2:end))/365.25;
```

   Compute the main descriptive statistics as the autocorrelation function and the periodogram and comment them.

   ii) Estimate the following model for these data:

$$
\begin{aligned}
y_t &= s_t + x_t \\
s_t &= \alpha_t \cos(\lambda t) + \alpha_t^* \sin(\lambda t) \\
(1-L)^d \alpha_t &= \eta_t \\
(1-L)^d \alpha_t^* &= \eta_t^* \\
(1-\phi L) x_t &= \varepsilon_t
\end{aligned}
\tag{2}
$$

   with $\eta_t, \eta_t^* \sim i.i.d.N(0, \sigma_\eta^2)$ and $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$. The frequency parameter λ is considered as known and estimated by the maximum of the periodogram. Provides the estimates of the remaining parameters along with relative comments and the plots of the residuals ACF and periodogram.

   iii) Extract the red noise component x_t. By comparing it with the plot of the US natural gas and oil production in figure 2, what are your comments?

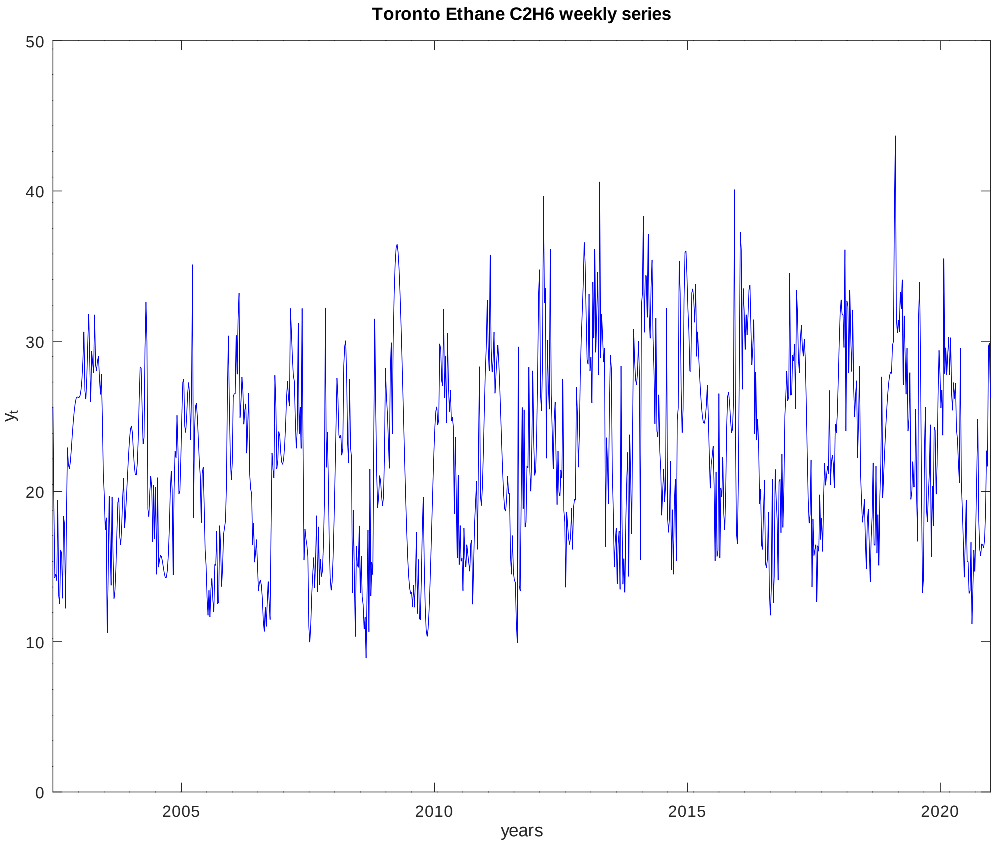
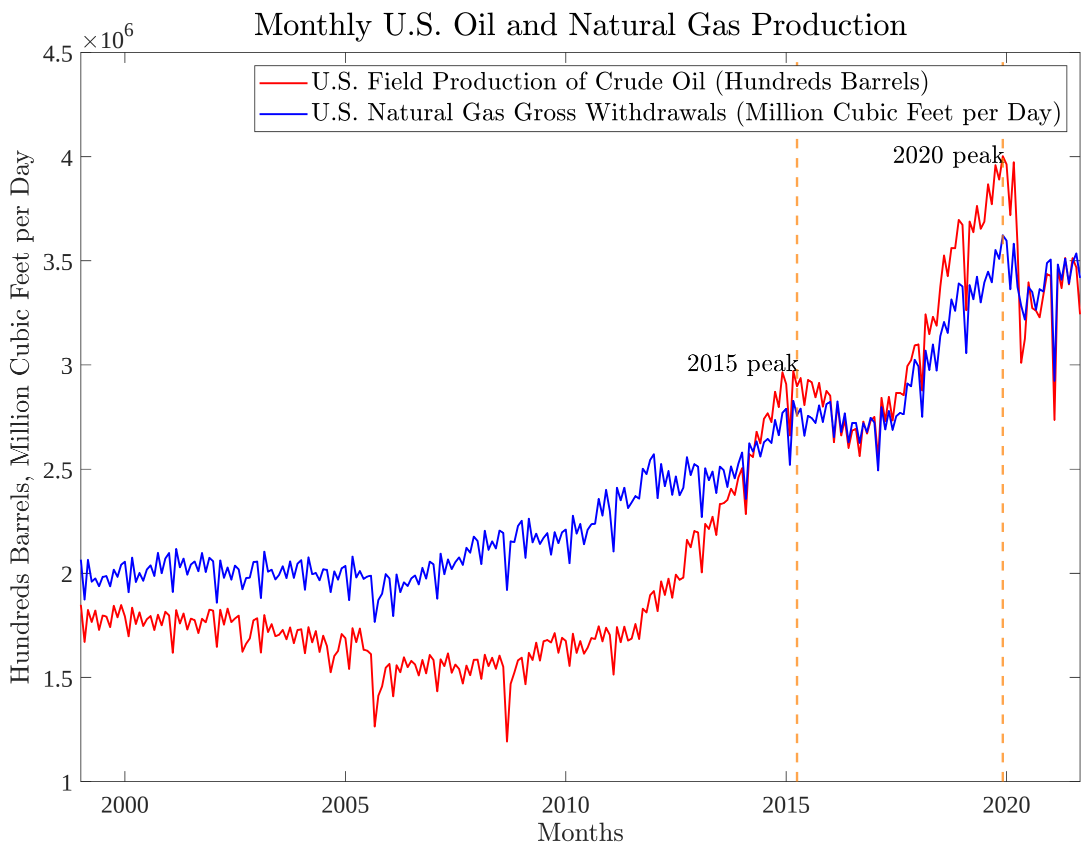


Figure 1: Weakly mean of the FTIR measurements of ethane abundance in the atmosphere recorded at Toronto (Canada).

Figure 2: Monthly United States natural gas production gross withdrawals (million cubic feet per day, blue line) and field production of crude oil (hundred barrels, red line). Source: U.S. Energy Information Administration, https://www.eia.gov.



References

Franco, B., Bader, W., Toon, G., Bray, C., Perrin, A., Fischer, E., ... Mahieu, E. (2015). Retrieval of ethane from ground-based ftir solar spectra using improved spectroscopy: Recent burden increase above jungfraujoch. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *160*, 36-49.

Franco, B., Mahieu, E., Emmons, L. K., Tzompa-Sosa, Z. A., Fischer, E. V., Sudo, K., ... Walker, K. A. (2016, apr). Evaluating ethane and methane emissions associated with the development of oil and natural gas extraction in north america. *Environmental Research Letters*, *11*(4), 044010.

Helmig, D., Rossabi, S., Hueber, J., and et al. (2016). Reversal of global atmospheric ethane and propane trends largely due to us oil and natural gas production. *Nature Geosciences*, *11*(9), 490-495.