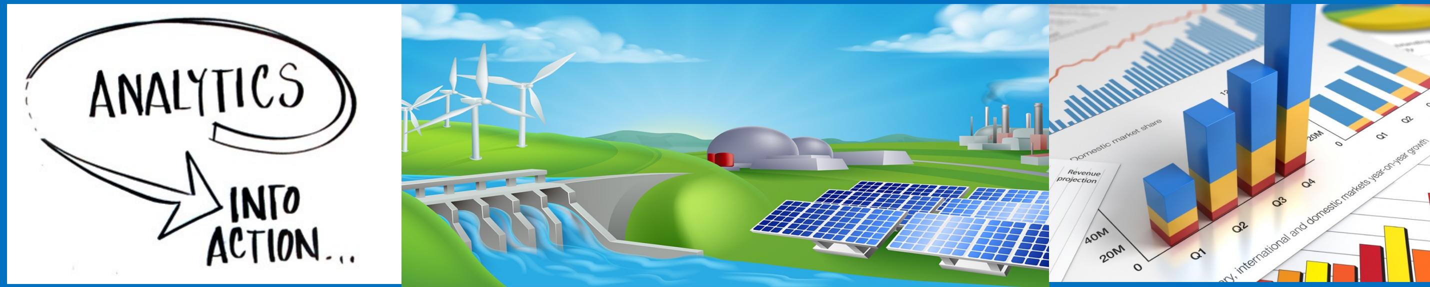



# TIME SERIES ANALYSIS FOR ENERGY DATA

# M8.1 – ARIMA Model Identification and Parameter Estimation

Prof. Luana Medeiros Marangon Lima, Ph.D.

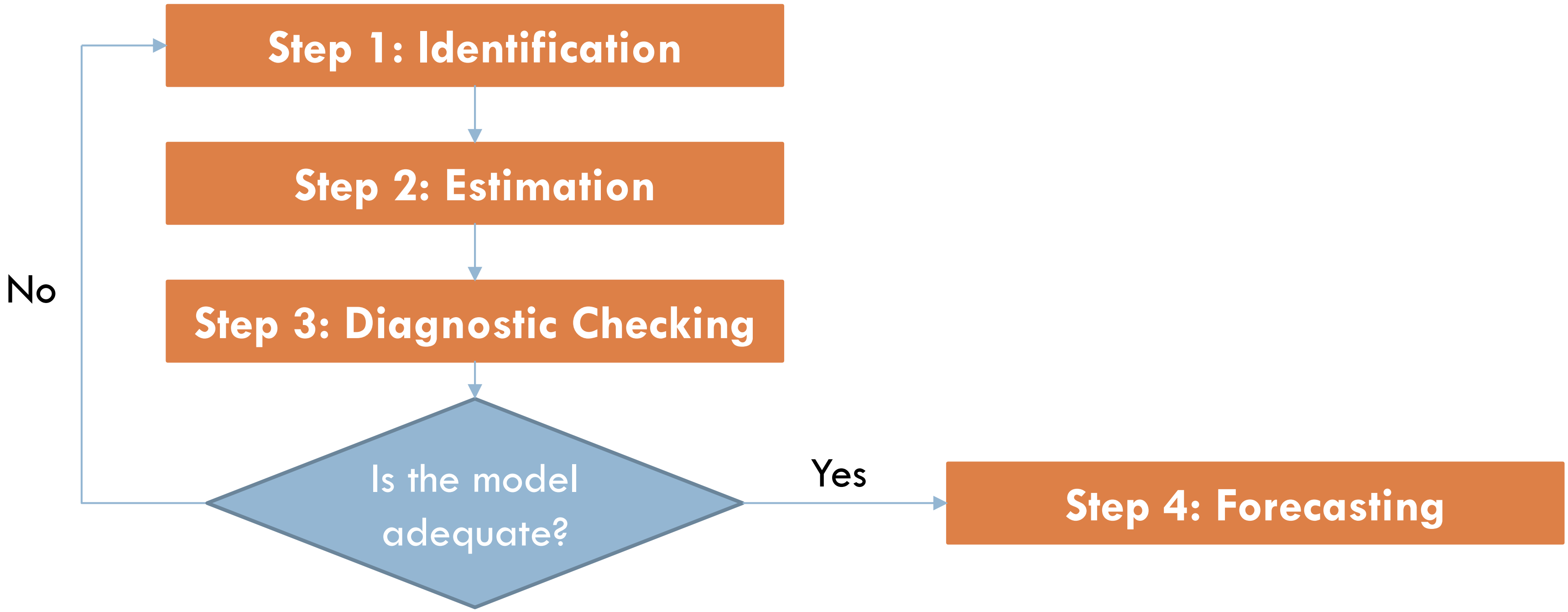
Nicholas School of the Environment - Duke University

# Learning Goals

- Simple Example of ACF and PACF computation
- ARIMA Model Parameters Estimation
  - AR coefficients
  - MA coefficients
  - Variance of residuals

# ARIMA Parameter Estimation

# ARIMA Modeling - Process



# Parameter Estimation

- We know that AR coefficients come from PACF
- How can we get PACF?
- Simple example: Excel spreadsheet
  - "*Temp_example_ACF_PACF_computation.xlsx*"

# Simple Example: ACF Computation

| $t$ | $Y_t$ |
|---|---|
| 1 | 14.2 |
| 2 | 16.4 |
| 3 | 11.9 |
| 4 | 15.2 |
| 5 | 18.5 |
| 6 | 22.1 |
| 7 | 19.4 |
| 8 | 25.1 |
| 9 | 23.4 |
| 10 | 18.1 |
| 11 | 22.6 |
| 12 | 17.2 |

**ACF(1)**

| | | Lag 1 | | | |
|---|---|---|---|---|---|
| t | Y_t | Y1_t | [Y_t - Mu(Y)] | [Y1_t - Mu(Y)] | Mult |
| 1 | 14,2 | - | - | - | - |
| 2 | 16,4 | 14,2 | -2,275 | -4,475 | 10,181 |
| 3 | 11,9 | 16,4 | -6,775 | -2,275 | 15,413 |
| 4 | 15,2 | 11,9 | -3,475 | -6,775 | 23,543 |
| 5 | 18,5 | 15,2 | -0,175 | -3,475 | 0,608 |
| 6 | 22,1 | 18,5 | 3,425 | -0,175 | -0,599 |
| 7 | 19,4 | 22,1 | 0,725 | 3,425 | 2,483 |
| 8 | 25,1 | 19,4 | 6,425 | 0,725 | 4,658 |
| 9 | 23,4 | 25,1 | 4,725 | 6,425 | 30,358 |
| 10 | 18,1 | 23,4 | -0,575 | 4,725 | -2,717 |
| 11 | 22,6 | 18,1 | 3,925 | -0,575 | -2,257 |
| 12 | 17,2 | 22,6 | -1,475 | 3,925 | -5,789 |
| **Mean** | **18,68** | | | **SUM** | **75,882** |
| **Std. Dev.** | **3,84** | | | **Num. Obs.** | **12** |
| | | | | **COVARIANCE(Y,Y1)** | **6,32** |
| | | | | **CORRELATION(Y,Y1)** | **0,4288** |

**ACF(2)**

| | | Lag 1 | | | |
|---|---|---|---|---|---|
| t | Y_t | Y2_t | [Y_t - Mu(Y)] | [Y2_t - Mu(Y)] | Mult |
| 1 | 14,2 | - | - | - | - |
| 2 | 16,4 | - | - | - | - |
| 3 | 11,9 | 14,2 | -6,775 | -4,475 | 30,318 |
| 4 | 15,2 | 16,4 | -3,475 | -2,275 | 7,906 |
| 5 | 18,5 | 11,9 | -0,175 | -6,775 | 1,186 |
| 6 | 22,1 | 15,2 | 3,425 | -3,475 | -11,902 |
| 7 | 19,4 | 18,5 | 0,725 | -0,175 | -0,127 |
| 8 | 25,1 | 22,1 | 6,425 | 3,425 | 22,006 |
| 9 | 23,4 | 19,4 | 4,725 | 0,725 | 3,426 |
| 10 | 18,1 | 25,1 | -0,575 | 6,425 | -3,694 |
| 11 | 22,6 | 23,4 | 3,925 | 4,725 | 18,546 |
| 12 | 17,2 | 18,1 | -1,475 | -0,575 | 0,848 |
| **Mean** | **18,68** | | | **SUM** | **68,511** |
| **Std. Dev.** | **3,84** | | | **Num. Obs.** | **12** |
| | | | | **CORRELATION(Y,Y2)** | **5,71** |
| | | | | **CORRELATION(Y,Y2)** | **0,3871** |

$$Cov(Y_t Y_s) = E[(Y_t - \mu)(Y_s - \mu)]$$

$$Cov(Y_t Y_s) = \frac{1}{n}\sum(Y_t - \mu)(Y_s - \mu)$$

$$Corr(Y_t Y_s) = \frac{Cov(Y_t Y_s)}{\sigma_Y^2} = \text{ACF(t−s}$$

Compare with R function:

```
> print(acf_temp$acf)
, , 1

           [,1]
 [1,]  1.00000000
 [2,]  0.42875355
 [3,]  0.38710748
 [4,]  0.13060966
 [5,] -0.24677581
 [6,] -0.36416383
 [7,] -0.30293249
 [8,] -0.27678937
 [9,] -0.23337053
[10,]  0.02054935
[11,] -0.08028336
```

# Simple Example: ACF Computation

| t | $Y_t$ |
|---|---|
| 1 | 14.2 |
| 2 | 16.4 |
| 3 | 11.9 |
| 4 | 15.2 |
| 5 | 18.5 |
| 6 | 22.1 |
| 7 | 19.4 |
| 8 | 25.1 |
| 9 | 23.4 |
| 10 | 18.1 |
| 11 | 22.6 |
| 12 | 17.2 |

**ACF(3)**

| | | Lag 1 | | | |
|---|---|---|---|---|---|
| t | Y_t | Y3_t | [Y_t - Mu(Y)] | [Y3_t - Mu(Y)] | Mult |
| 1 | 14,2 | - | - | - | - |
| 2 | 16,4 | - | - | - | - |
| 3 | 11,9 | - | - | - | - |
| 4 | 15,2 | 14,2 | -3,475 | -4,475 | 15,551 |
| 5 | 18,5 | 16,4 | -0,175 | -2,275 | 0,398 |
| 6 | 22,1 | 11,9 | 3,425 | -6,775 | -23,204 |
| 7 | 19,4 | 15,2 | 0,725 | -3,475 | -2,519 |
| 8 | 25,1 | 18,5 | 6,425 | -0,175 | -1,124 |
| 9 | 23,4 | 22,1 | 4,725 | 3,425 | 16,183 |
| 10 | 18,1 | 19,4 | -0,575 | 0,725 | -0,417 |
| 11 | 22,6 | 25,1 | 3,925 | 6,425 | 25,218 |
| 12 | 17,2 | 23,4 | -1,475 | 4,725 | -6,969 |
| Mean | 18,68 | | | SUM | 23,116 |
| Std. Dev. | 3,84 | | | Num. Obs. | 12 |
| | | | | CORRELATION(Y,Y2) | 1,93 |
| | | | | CORRELATION(Y,Y2) | 0,1306 |

**ACF(4)**

| | | Lag 1 | | | |
|---|---|---|---|---|---|
| t | Y_t | Y4_t | [Y_t - Mu(Y)] | [Y4_t - Mu(Y)] | Mult |
| 1 | 14,2 | - | - | - | - |
| 2 | 16,4 | - | - | - | - |
| 3 | 11,9 | - | - | - | - |
| 4 | 15,2 | - | - | - | - |
| 5 | 18,5 | 14,2 | -0,175 | -4,475 | 0,783 |
| 6 | 22,1 | 16,4 | 3,425 | -2,275 | -7,792 |
| 7 | 19,4 | 11,9 | 0,725 | -6,775 | -4,912 |
| 8 | 25,1 | 15,2 | 6,425 | -3,475 | -22,327 |
| 9 | 23,4 | 18,5 | 4,725 | -0,175 | -0,827 |
| 10 | 18,1 | 22,1 | -0,575 | 3,425 | -1,969 |
| 11 | 22,6 | 19,4 | 3,925 | 0,725 | 2,846 |
| 12 | 17,2 | 25,1 | -1,475 | 6,425 | -9,477 |
| Mean | 18,68 | | | SUM | -43,675 |
| Std. Dev. | 3,84 | | | Num. Obs. | 12 |
| | | | | CORRELATION(Y,Y2) | -3,64 |
| | | | | CORRELATION(Y,Y2) | -0,2468 |

Compare with R values:

```
> print(acf_temp$acf)
, , 1

          [,1]
 [1,]  1.00000000
 [2,]  0.42875355
 [3,]  0.38710748
 [4,]  0.13060966
 [5,] -0.24677581
 [6,] -0.36416383
 [7,] -0.30293249
 [8,] -0.27678937
 [9,] -0.23337053
[10,]  0.02054935
[11,] -0.08028336
```

# PACF Concept

- Consider the two regression models

$$\boxed{y_t = \beta_0 + \beta_1 y_{t-2}}$$

$$\boxed{y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2}}$$

- What is the meaning of $\beta_1$ in the first model and $\beta_2$ in the second model?
- They both represent the linear dependence between observation $y_t$ and $y_{t-2}$
- But what's difference between $\beta_1$ in the first model and $\beta_2$ in the second model?
- **The $\boldsymbol{\beta_2}$ in the second model is the linear dependence between $\boldsymbol{y_t}$ and $\boldsymbol{y_{t-2}}$ WITH the dependency between $\boldsymbol{y_t}$ and $\boldsymbol{y_{t-1}}$ ALREADY accounted for**

# ARIMA Parameter Estimation

- What do we need to estimate?
  - AR coefficients
  - MA coefficients
  - Variance of the residuals (error or innovation)
- Estimation Methods for ARIMA coefficients
  - Least squares method
  - Maximum Likelihood
  - Methods of Moments or Yule-Walker equations
- Other Methods: Bayesian estimation or Kalman Filtering

# Estimating AR(p) parameters

- AR(p)

$$Y_t = \phi_1 \ddot{Y}_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t \qquad for\ t = 1,2,\ldots,n$$
$$a_t \sim i.i.d.(0, \sigma^2)$$

- Need to estimate $\phi = \left(\phi_1\ \phi_2\ \ldots \phi_p\right)'$ and $\sigma^2$
- Estimation method: Method of moments (Yule-Walker equations)
- Yule Walker equations relate AR model coefficients to the autocovariance (ACF) of the random process
- Note: They do not address the model order, they are simply used to estimate AR parameters

# Estimating AR coefficients

- Consider AR(p) model with a zero mean

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t$$

$$Y_t Y_{t-h} = \phi_1 Y_{t-1} Y_{t-h} + \phi_2 Y_{t-2} Y_{t-h} + \cdots + \phi_p Y_{t-p} Y_{t-h} + a_t\, Y_{t-h} \qquad (\boldsymbol{with}\ \boldsymbol{h} > \mathbf{0})$$

$$E[Y_t Y_{t-h}] = E[\phi_1 Y_{t-1} Y_{t-h}] + E[\phi_2 Y_{t-2} Y_{t-h}] + \cdots + E[\phi_p Y_{t-p} Y_{t-h}] + E[a_t\, Y_{t-h}]$$

$$E[Y_t Y_{t-h}] = \phi_1 E[Y_{t-1} Y_{t-h}] + \cdots + \phi_p E[Y_{t-p} Y_{t-h}] + E[a_t Y_{t-h}] \qquad \boldsymbol{Eq}.(\mathbf{1})$$

- Closer look at last term $E[a_t Y_{t-h}]$
  - $a_t$ is uncorrelated with $Y_{t-h}$, therefore $E[a_t Y_{t-h}] = \mathrm{a_t} E[Y_{t-h}]$
  - $E[Y_t] = 0$ since this is a zero mean process, therefore

$$E[a_t Y_{t-h}] = \mathrm{a_t}\ E[Y_{t-h}] = \mathrm{a_t} * 0 = 0$$

- Back in Eq. (1)

$$E[Y_t Y_{t-h}] = \phi_1 E[Y_{t-1} Y_{t-h}] + \cdots + \phi_p E[Y_{t-p} Y_{t-h}] + \underbrace{E[a_t Y_{t-h}]}_{\mathbf{0}}$$

$$E[Y_t Y_{t-h}] = \phi_1 E[Y_{t-1} Y_{t-h}] + \cdots + \phi_p E[Y_{t-p} Y_{t-h}] \qquad \boldsymbol{Eq}.(\mathbf{2})$$

# Estimating AR coefficients (cont'd)

- Recall two definitions from Lect. 3
  1. $\gamma_{t,s} = Cov(Y_t Y_s) = E[Y_t Y_s] - \mu_t \mu_s$
  2. For a stationary process $\gamma_{t,s} = \gamma_{0,|t-s|}$
- Since we are considering a zero mean series the above relationships leads to

$\gamma_{t,s} = E[Y_t Y_s] - \overset{0}{\mu_t \mu_s} \therefore E[Y_t Y_s] = \gamma_{t,s}$

- Rewriting in terms of $t$ and $t - h$ we get

$E[Y_t Y_{t-h}] = \gamma_{t,t-h} = \gamma_{0,h}$ or simply $E[Y_t Y_{t-h}] = \gamma_h$

- Substituting in Eq. (2)

$E[Y_t Y_{t-h}] = \phi_1 E[Y_{t-1} Y_{t-h}] + \cdots + \phi_p E\left[Y_{t-p} Y_{t-h}\right]$

$\gamma_h = \phi_1 \gamma_{h-1} + \cdots + \phi_p \gamma_{h-p}$

# Estimating AR coefficients (cont'd)

- For a zero mean process autocovariance divided by variance equal autocorrelation, therefore

$$\rho_h = \phi_1 \rho_{h-1} + \cdots + \phi_p \rho_{h-p}$$

- Writing this equation for $h = 1,2,\dots,p$ we get

Note that for h=1:
$\rho_{h-p} = \rho_{1-p} = \rho_{p-1}$

$$\left.\begin{array}{ll} \boldsymbol{h=1} & \rho_1 = \phi_1 \rho_0 + \phi_2 \rho_1 + \cdots + \boxed{\phi_p \rho_{p-1}} \\ \boldsymbol{h=2} & \rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 + \cdots + \phi_p \rho_{p-2} \\ \vdots & \vdots \\ \boldsymbol{h=p} & \rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \cdots + \phi_p \rho_0 \end{array}\right\}$$

***Yule-Walker equations***

# Estimating AR coefficients (cont'd)

- In matrix form

$$\underbrace{\begin{bmatrix}\rho_1\\ \rho_2\\ \rho_3\\ \vdots\\ \rho_p\end{bmatrix}}_{\boldsymbol{\rho}} = \underbrace{\begin{bmatrix}\rho_0 & \rho_1 & \rho_2 & \cdots & \rho_{p-1}\\ \rho_1 & \rho_0 & \rho_1 & \cdots & \rho_{p-2}\\ \rho_2 & \rho_1 & \rho_0 & \cdots & \rho_{p-3}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \rho_0\end{bmatrix}}_{\boldsymbol{R}} \cdot \underbrace{\begin{bmatrix}\phi_1\\ \phi_2\\ \phi_3\\ \vdots\\ \phi_p\end{bmatrix}}_{\boldsymbol{\phi}}$$

- This a linear system with $p$ equations and $p$ variables

$$\boldsymbol{\rho_p} = \boldsymbol{R_p} \cdot \boldsymbol{\phi_p} \quad \Rightarrow \quad \boldsymbol{\phi_\mathrm{p}} = \boldsymbol{R_p^{-1}} \cdot \boldsymbol{\rho_p}$$

# Simple Example: PACF Computation

**PACF(2)**

**ACFs (\rho)**

| | |
|---|---|
| 0,4288 | |
| 0,3871 | |

=

**R (ACF matrix)**

| | |
|---|---|
| 1 | 0,4288 |
| 0,4288 | 1 |

*

| |
|---|
| \phi_1 |
| \phi_2 |

| |
|---|
| \phi_1 |
| \phi_2 |

=

**R inverse**

| | |
|---|---|
| 1,2252 | -0,5253 |
| -0,5253 | 1,2252 |

*

**ACFs (\rho)**

| |
|---|
| 0,4288 |
| 0,3871 |

=

| |
|---|
| 0,322 |
| **0,249** |

Compare with R values:

```
> print(pacf_temp$acf)
, , 1

          [,1]
 [1,]  0.4287535
 [2,]  0.2490630
 [3,] -0.1316882
 [4,] -0.4645661
 [5,] -0.2730998
 [6,]  0.2045947
 [7,]  0.1920660
 [8,] -0.2915612
 [9,] -0.1206727
[10,] -0.1351969
```

# Estimating the variance of AR(p)

- To get an estimate for the variance use same approach

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t$$

- Multiply by $Y_{t-h}$ and take expectations

$$E[Y_t Y_{t-h}] = \phi_1 E[Y_{t-1} Y_{t-h}] + \cdots + \phi_p E\left[Y_{t-p} Y_{t-h}\right] + E[a_t Y_{t-h}]$$

- But now we consider the case where $h = 0$, so the last term is no equal to zero, instead

$$E[a_t Y_t] = \sigma^2$$

- Therefore

$$\rho_0 = \phi_1 \rho_1 + \cdots + \phi_p \rho_p + \sigma^2 \Rightarrow \sigma^2 = \rho_0 - \phi_1 \rho_1 - \cdots - \phi_p \rho_p$$

- From estimates of $\phi_1, \ldots, \phi_p$ we can get an estimate for $\sigma^2$

# Estimating MA coefficients

- Difficult because regressors are unknown residuals
- Assuming **invertibility** we can right a MA process as an AR and then use the Yule-Walker equations
- Example: MA(1) to AR(∞)
  - MA(1): $Y_t = \epsilon_t - \theta\epsilon_{t-1}$
  - Define operator L such that: $L\epsilon_t = \epsilon_{t-1}$ (also denoted by B and know as back shift operator)
    - Therefore $LL\epsilon_t = L\epsilon_{t-1} = \epsilon_{t-2}$
  - In terms of back shift operator: $Y_t = (1 - \theta L)\epsilon_t$
  - Rewriting: $\frac{Y_t}{(1-\theta L)} = \epsilon_t$

Sum of Geometric Progression: $S_\infty = \mathrm{a} + \mathrm{ra} + \mathrm{r}^2\mathrm{a} + \cdots = \frac{\mathrm{a}}{1-r}$ if $|r| < 1$

# Estimating MA coefficients (cont'd)

- In that case if $|\theta| < 1$

$$\frac{Y_t}{(1-\theta L)} = Y_t + \theta L Y_t + \theta^2 L^2 Y_t + \theta^3 L^3 Y_t + \cdots = \epsilon_t$$

- Isolating $Y_t$ we get

$$Y_t = -\theta \underbrace{LY_t}_{Y_{t-1}} - \theta^2 \underbrace{L^2 Y_t}_{Y_{t-2}} - \theta^3 \underbrace{L^3 Y_t}_{Y_{t-3}} - \cdots + \epsilon_t$$

From the definition of operator L

- Therefore

$$Y_t = -\theta Y_{t-1} - \theta^2 Y_{t-2} - \theta^3 Y_{t-3} - \cdots + \epsilon_t$$

Now you can use the same AR methods

# Example of Parameter Estimation

- Temperature Data

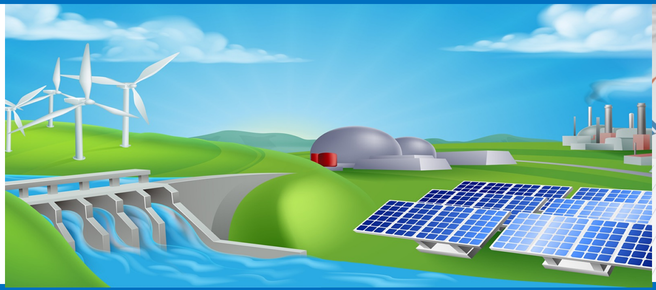
```
Series: temp
ARIMA(2,0,0) with non-zero mean

Coefficients:
         ar1     ar2     mean
      0.3190  0.2711  -0.4614
s.e.  0.2803  0.2907   2.0245

sigma^2 estimated as 14.31:  log likelihood=-31.45
AIC=70.9   AICc=76.62   BIC=72.84

Call:
lm(formula = Y[, 1] ~ Y[, 2] + Y[, 3])

Coefficients:
(Intercept)       Y[, 2]       Y[, 3]
     0.6157       0.2317       0.3206

[1] "Yule Walker results:"
          [,1]     [,2]
[1,] 0.3219669 0.249063
```

$Y_{t-1}$ $Y_{t-2}$

- Electricity Prices Data

```
Series: random_price
ARIMA(2,0,0) with non-zero mean

Coefficients:
          ar1     ar2    mean
      -0.0711  0.0335  0.0000
s.e.   0.0706  0.0710  0.0059

sigma^2 estimated as 0.007577:  log likelihood=208.03
AIC=-408.06   AICc=-407.85   BIC=-394.82

Call:
lm(formula = Y[, 1] ~ Y[, 2] + Y[, 3])

Coefficients:
(Intercept)       Y[, 2]       Y[, 3]
  -0.001259     0.045078     0.043987

[1] "Yule Walker results:"
             [,1]       [,2]
[1,] -0.07055107 0.03261288
```

$Y_{t-1}$ $Y_{t-2}$




# THANK YOU !

luana.marangon.lima@duke.edu