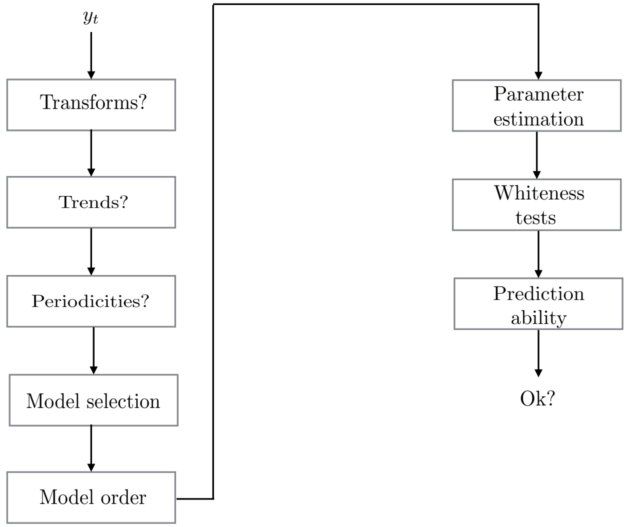
Time Series Analysis

Andreas Jakobsson

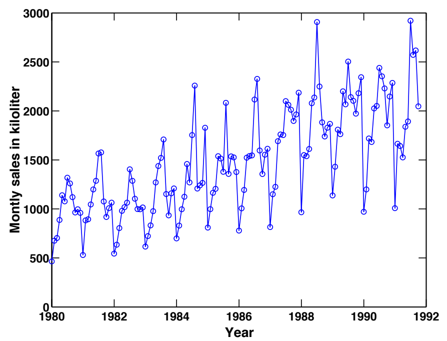
Content:

- Trends
- Transforms

The identification process



Trends



Some measurements contain trends or strong periodicities; these should typically be modelled separately, before selecting an AR, MA, or ARMA model.

Deterministic trends

Generally, one treats trends as either being deterministic or stochastic. In the former case, one assumes that the trend is some deterministic function that has been added to the (assumed) stationary time series; for example, the trend may be a polynomial trend

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_k t^k + x_t$$

where α_ℓ, for $\ell = 1, \ldots, k$, denote (unknown) deterministic constants and x_t a stochastic process. The trend may also be cyclic, such that

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 \cos(\omega t + \theta) + x_t \\ &= \alpha_0 + \beta_1 \cos(\omega t) + \gamma_1 \sin(\omega t) + x_t \end{aligned}$$

where ω and θ denote the unknown frequency and phase of the periodicity, and with $\beta_1 = \alpha_1 \cos(\theta)$ and $\gamma_1 = -\alpha_1 \sin(\theta)$. More generally, the model can consist of several periodicities, such that

$$y_t = \alpha_0 + \sum_{\ell=1}^{d} \beta_\ell \cos(\omega_\ell t) + \gamma_\ell \sin(\omega_\ell t) + x_t$$

Stochastic trends

To handle stochastic trends, one often instead use an ARIMA or SARIMA model. An autoregressive integrated moving average (ARIMA) process of order (p,d,q) is defined as

$$A(z)(1 - z^{-1})^d y_t = C(z)e_t$$

where d denotes the number of differentiations, typically being $d = 0, 1,$ or 2.

Stochastic trends

To handle stochastic trends, one often instead use an ARIMA or SARIMA model. An autoregressive integrated moving average (ARIMA) process of order (p,d,q) is defined as

$$A(z)(1 - z^{-1})^d y_t = C(z)e_t$$

where d denotes the number of differentiations, typically being $d = 0, 1,$ or 2.

A seasonal ARIMA (SARIMA) process is defined as

$$A(z)\mathcal{A}(z^s)\nabla^d\nabla_s^D y_t = C(z)\mathcal{C}(z^s)e_t$$

where

$$\nabla y_t = (1 - z^{-1})y_t$$

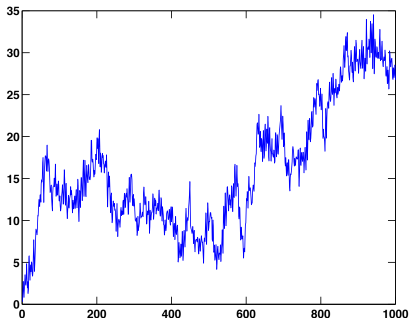
denotes the differentiation operator; ∇^d thus implies that the process is differentiated (recursively) d times. Furthermore,

$$\nabla_s y_t = y_t - y_{t-s}$$

where s is the seasonal period, and

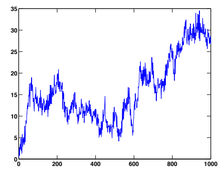
$$\mathcal{A}(z^s) = 1 + \breve{a}_1 z^{-s} + \ldots + \breve{a}_P z^{-sP}$$
$$\mathcal{C}(z^s) = 1 + \breve{c}_1 z^{-s} + \ldots + \breve{c}_Q z^{-sQ}$$

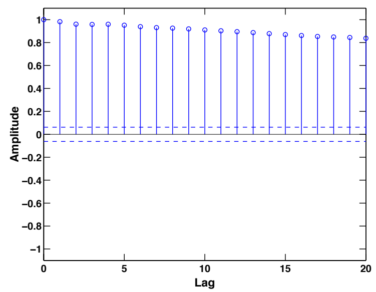
Trends



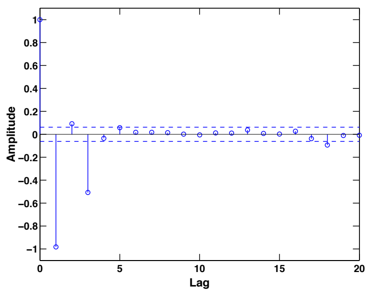
What would be a suitable model for this data?

Trends



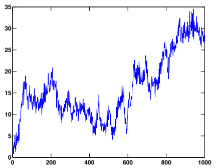


ACF

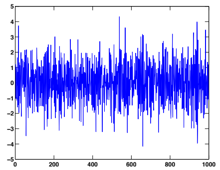


PACF

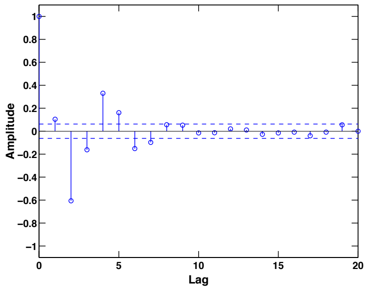
Trends



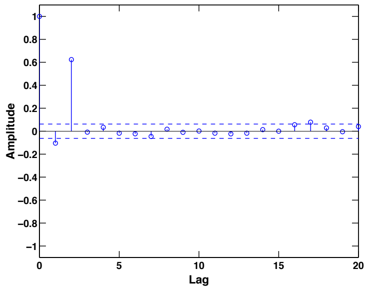
y_t



$\nabla y_t = (1 - z^{-1}) y_t$

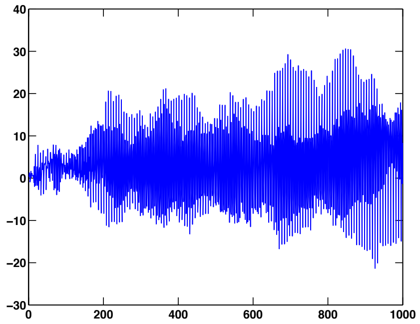


ACF



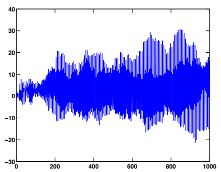
PACF

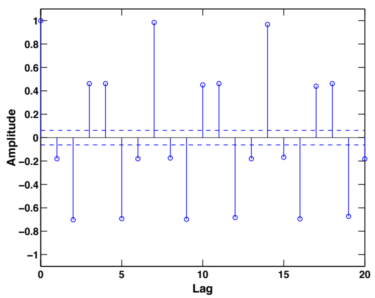
Trends



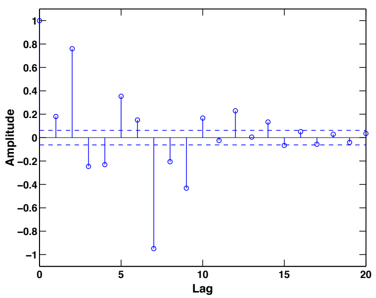
What would be a suitable model for this data?

Trends



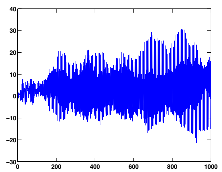


ACF

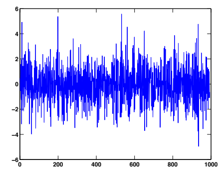


PACF

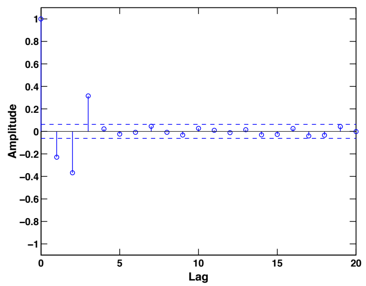
Trends



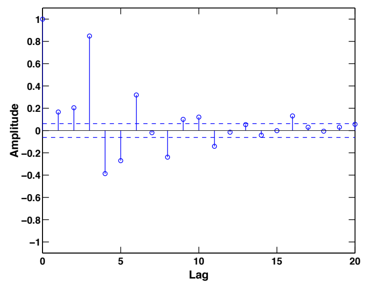
y_t



$\nabla_7 y_t = (1 - z^{-7}) y_t$

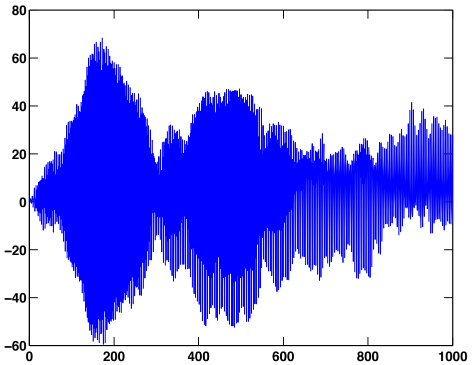


ACF



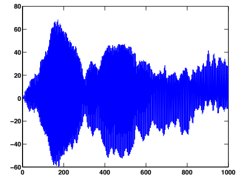
PACF

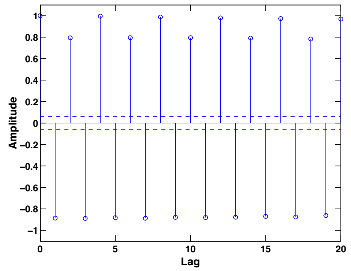
Trends



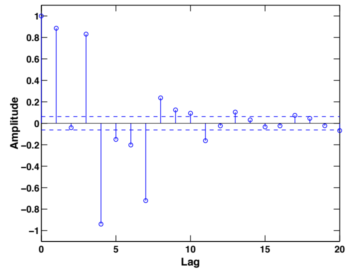
What would be a suitable model for this data?

Trends



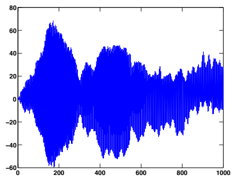


ACF

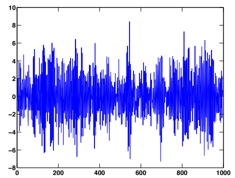


PACF

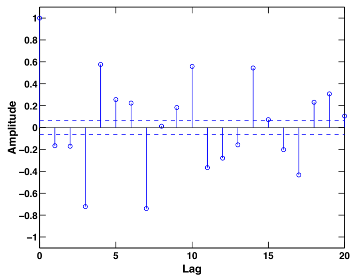
Trends



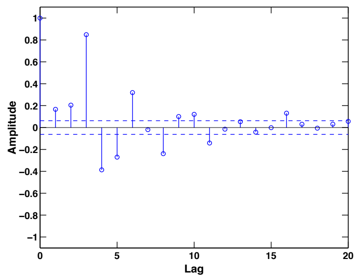
y_t



$\nabla_4 y_t = (1 - z^{-4}) y_t$

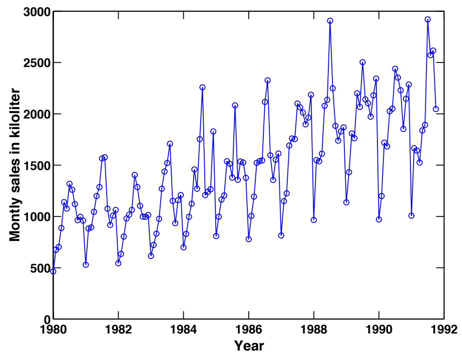


ACF



PACF

Transforms



Sometimes, one may get models by transforming the data.

Transforms

A suitable power transformation that may be used to stabilize the variance can be found as

$$y_t^{(\lambda)} = \begin{cases} \lambda^{-1}\left(y_t^{\lambda} - 1\right) & \lambda \neq 0 \\ \log(y_t) & \lambda = 0 \end{cases}$$

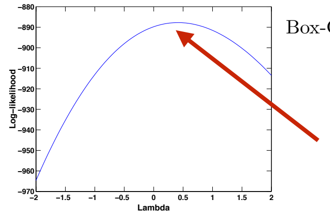
where λ should be selected to maximize the log-likelihood

$$L(\lambda) = -\frac{N}{2}\log\left\{\hat{\sigma}_y^2(\lambda)\right\} + (\lambda - 1)\sum_{t=1}^{N}\log(y_t)$$

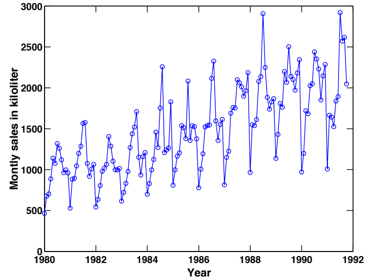
with $\hat{\sigma}_y(\lambda)$ denoting the estimated standard deviation of the transformed data, using the parameter λ.

Values of λ	-2.0	-1.0	-0.5	0.0	0.5	1.0	2.0
Transformation	y_t^{-2}	y_t^{-1}	$y_t^{-1/2}$	$\log(y_t)$	$\sqrt{y_t}$	y_t	y_t^2

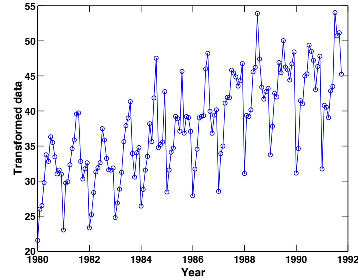
Transforms



Box-Cox plot

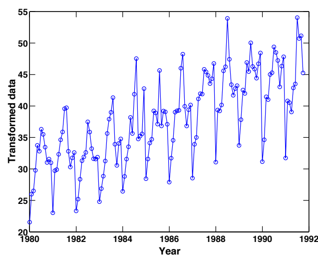


y_t

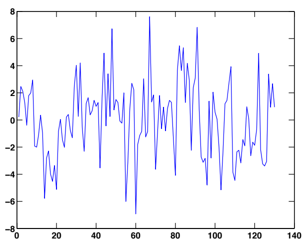


$\sqrt{y_t}$

Transforms



$\sqrt{y_t}$

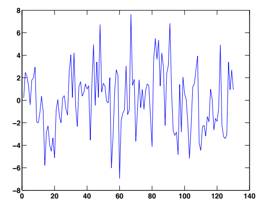


$\nabla_{12}\sqrt{y_t}$

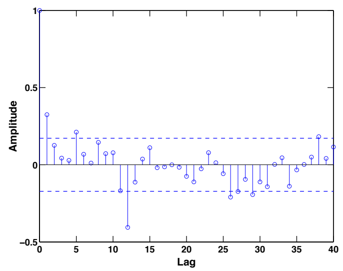
The differentiated time series does not have zero mean. We therefore form a new time series that will be zero-mean as

$$w_t = \nabla_{12}\sqrt{y_t} - 0.0681$$

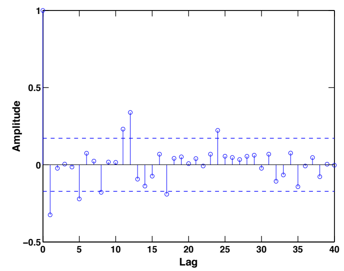
Transforms



$$w_t = \nabla_{12}\sqrt{y_t} - 0.0681$$



ACF



PACF