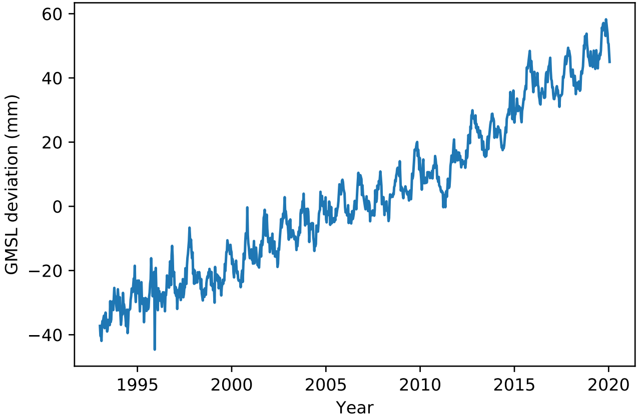
Time Series and Sequence Learning

Classical regression in a time series context

Fredrik Lindsten, Linköping University

ex) Global Mean Sea Level



What do you see in the data?

Data from **`https://climate.nasa.gov/vital-signs/sea-level/`**

Linear trend model

$$Y_t = \theta_0 + \theta_1 u_t + \underline{\varepsilon_t}$$ Error term

u_t = "time when observation #t was recorded"

$$u_t = 1993\frac{4}{365}, \ldots, 2020\frac{20}{365}$$

White noise

We model the errors as independent random variables

$$\varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2), \qquad\qquad t = 1,\, 2,\, \ldots$$

White noise

We model the errors as independent random variables

$$\varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2), \qquad\qquad t = 1, 2, \ldots$$

In a time series context, such a sequence of random variables is referred to as (Gaussian) **white noise**.

Linear trend model, cont'd

Simple linear trend model: $y_t = \theta_0 + \theta_1 u_t + \varepsilon_t$

The model parameters $\theta = (\theta_0, \theta_1)$ are estimated using OLS (= maximum likelihood)

Oridinary Least sequare

$$\hat{\theta} = \arg\min_{\theta} \sum_{t=1}^{n} (y_t - \{\theta_0 + \theta_1 u_t\})^2$$

$$\hat{\theta} = (\underline{\Phi}^T \underline{\Phi})^{-1} \underline{\Phi}^T \boldsymbol{y}$$

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad \underline{\Phi} = \begin{bmatrix} 1 & u_1 \\ \vdots & \vdots \\ 1 & u_n \end{bmatrix}$$

theta 0 is from the original formula
Theta1 need to calc the formula

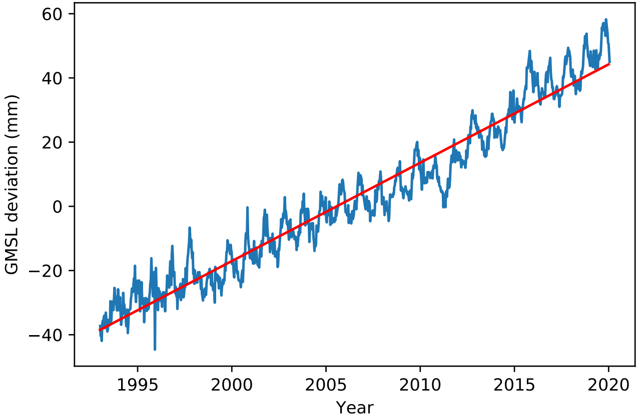
$$\hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{u}, \qquad \bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

cov

$$\hat{\theta}_1 = \frac{\sum_{t=1}^{n} (y_t - \bar{y})(u_t - \bar{u})}{\sum_{t=1}^{n} (u_t - \bar{u})^2}, \qquad \bar{u} = \frac{1}{n}\sum_{t=1}^{n} u_t$$

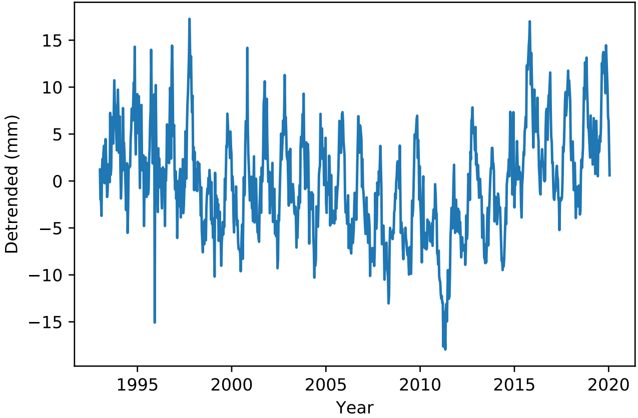
var $\longrightarrow$

ex) Global Mean Sea Level



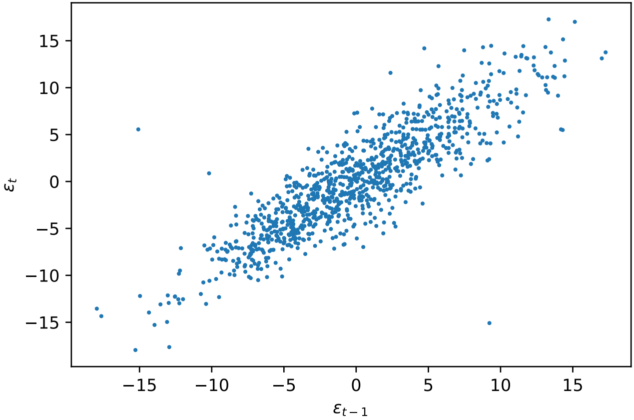
Linear trend fitted with OLS

ex) Global Mean Sea Level



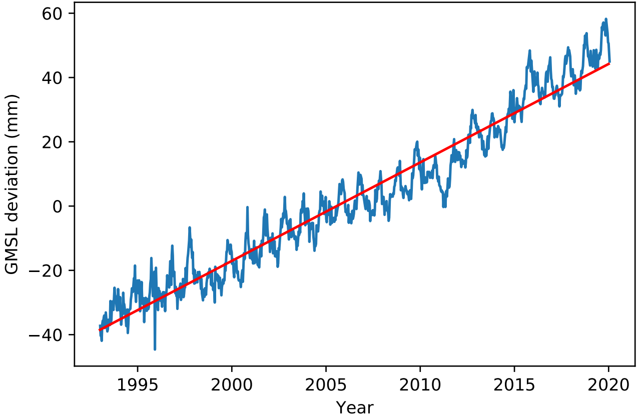
Residuals – Do they match the model?

ex) Global Mean Sea Level



Scatter plot of residuals at lag 1

ex) Global Mean Sea Level



Linear trend fitted with OLS

Classical regression in a time series context

Classical regression is insufficient!

Classical regression is insufficient!

Why?

- Time series data have temporal dependencies
- Forecasting = extrapolation

ex) *The number of infected individuals tomorrow is **statistically dependent** on the number of infected individuals today*