Time series analysis and forecasting

Integrative analytics for multi-modal, multi-scale neuroscience

IPN Summer School

Golia Shafiei, PhD candidate

Supervisor: Dr. Bratislav Misic

Network Neuroscience Lab

McConnell Brain Imaging Centre

Montreal Neurological Institute, McGill University

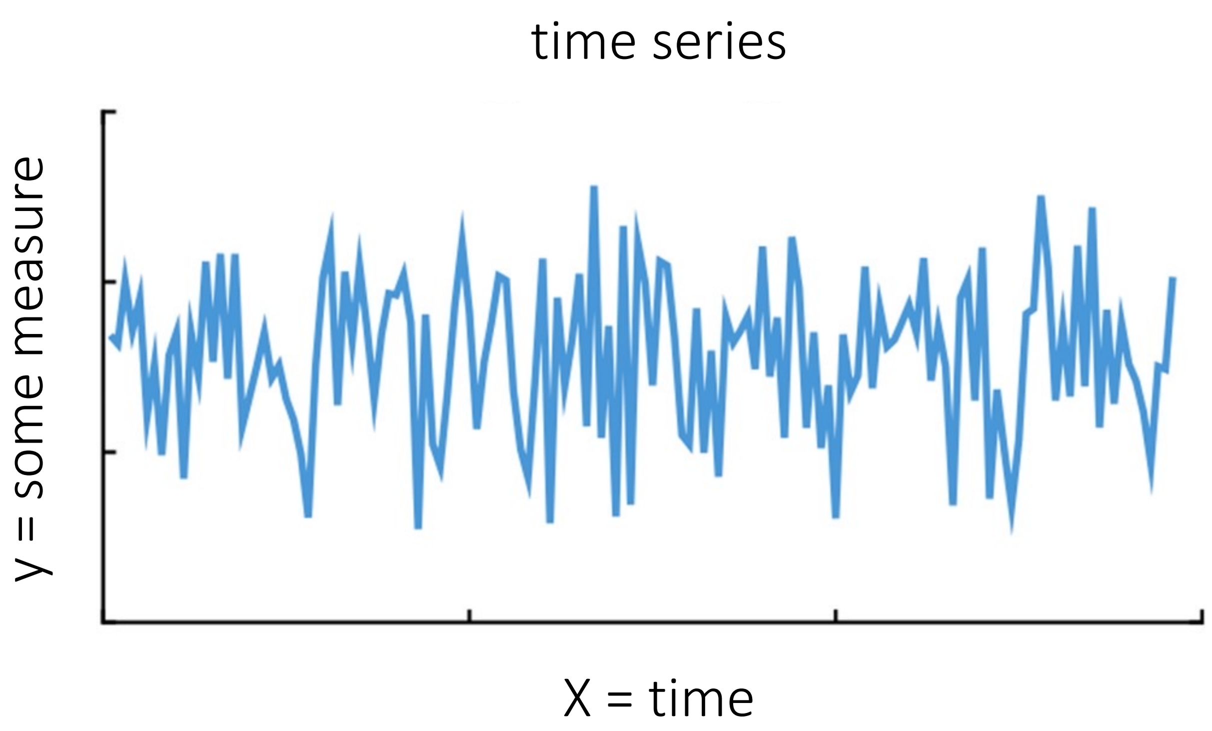
June 30, 2021

Content

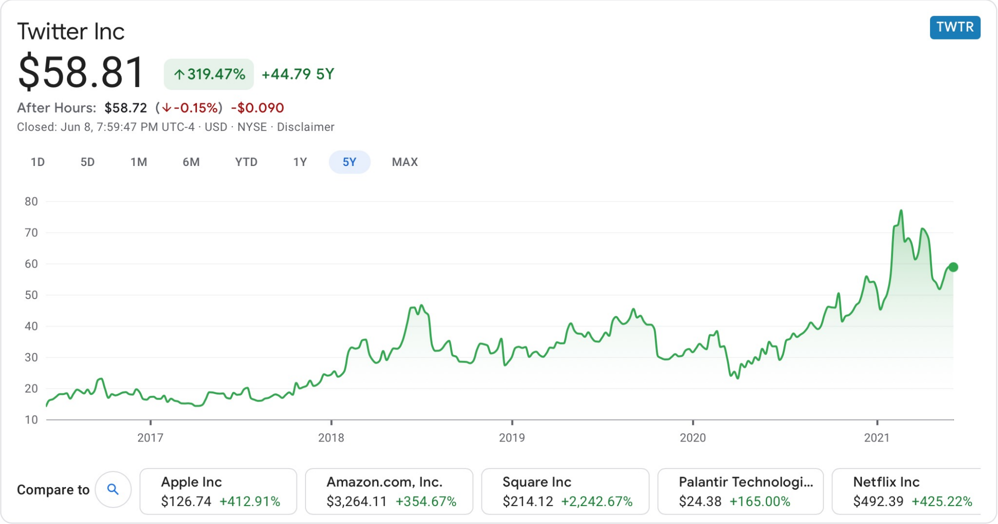
- Definition and characteristics of a time series
- Time series analysis
- Time series forecasting
- Spectral analysis
- Time series phenotyping

What is a time series?

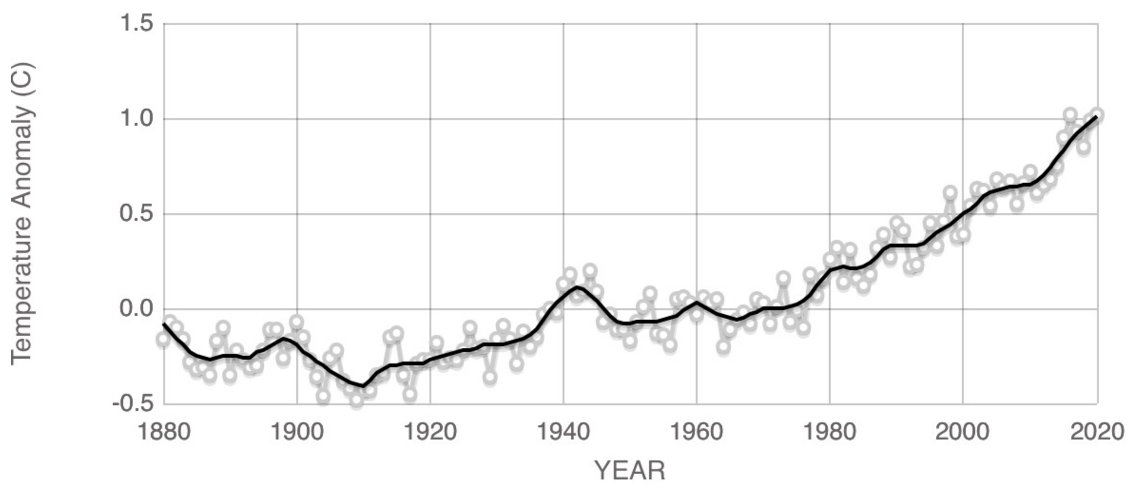
- Repeated measurements of a variable over time



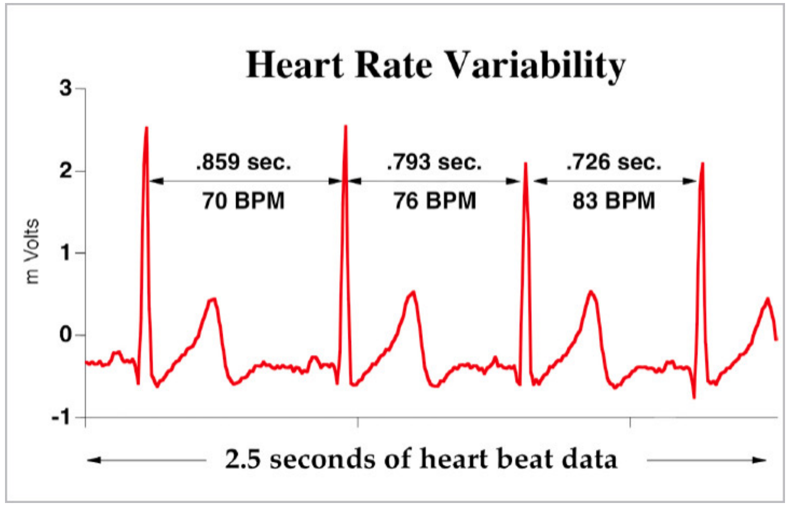
What is a time series?



Source: https://www.google.com/finance/

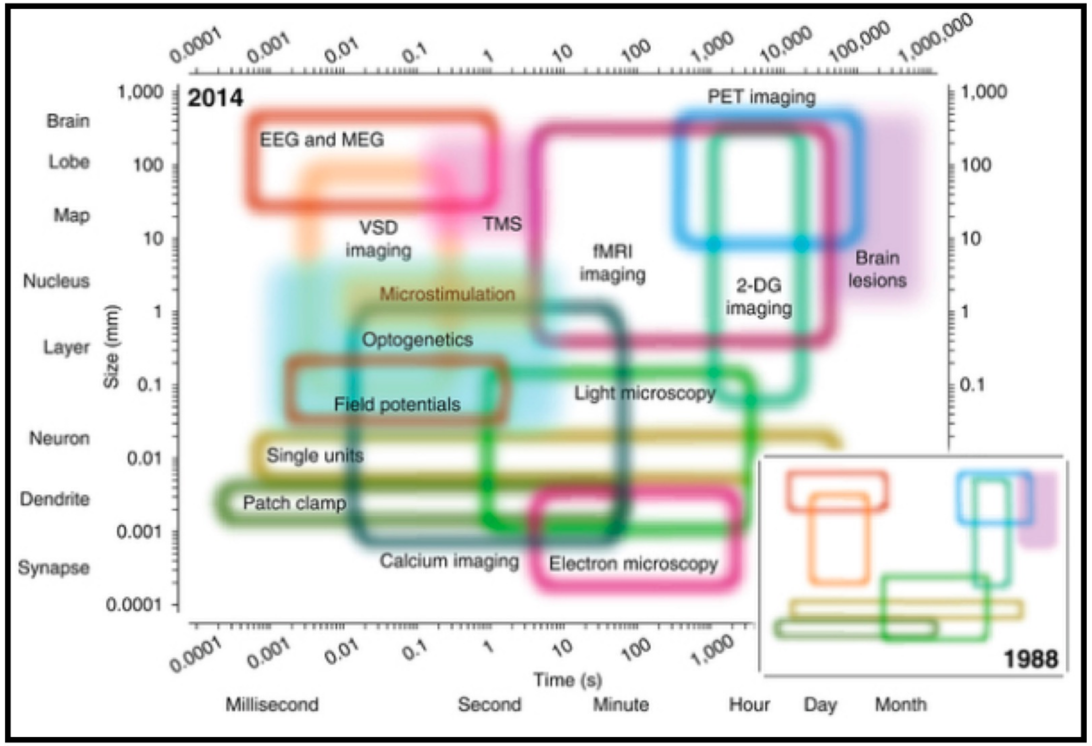


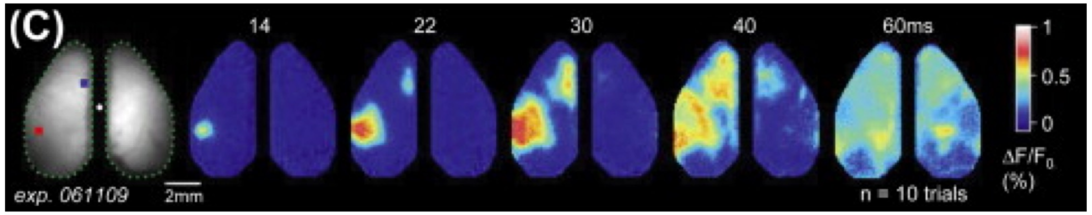
Source: https://climate.nasa.gov



What is a time series?

- Neural time series measured at different spatial and temporal scales





Electrophysiology:
> voltage [V(t)]
> currents (ion exchange)
> **MEG**: magnetic fields

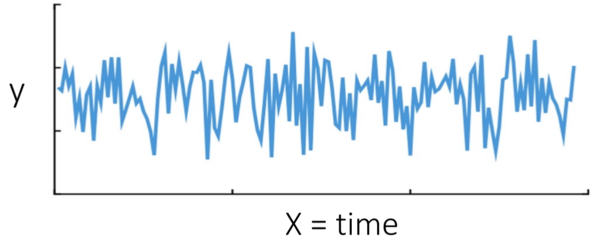
Optical Imaging: (luminance)
Concentration of something
> Calcium
> Non-specific charges
(voltage sensitive dye)

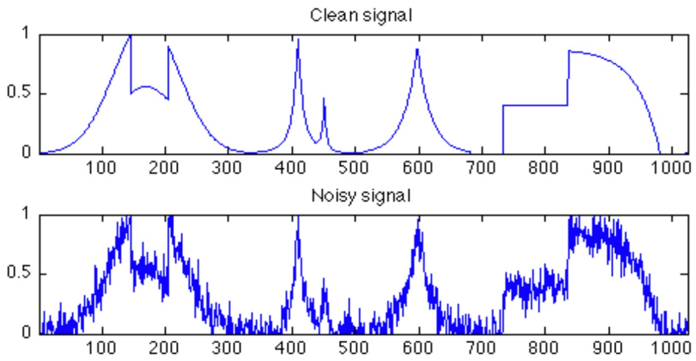
Medical Imaging:
> fMRI: magnetic spin
> PET: radiation (positron)

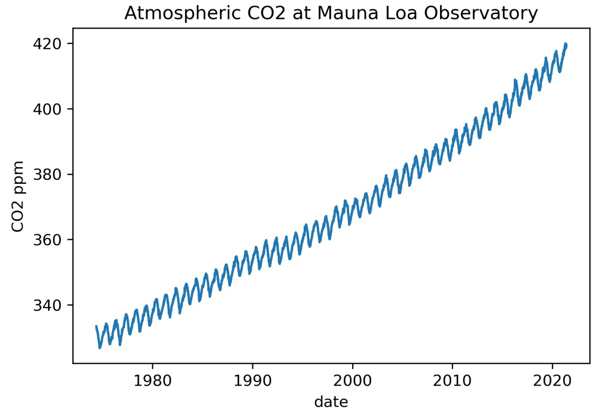
Neural Signal Processing course by Richard Gao: https://github.com/rdgao/COGS118C

Characteristics of time series data

- An ordered sequence of values of a variable measured over time
  - Time is a monotonically increasing axis
  - Time series vs randomly sampled data
- Structure and behavior of data
  - Signal and noise components
  - Stationarity vs non-stationarity
  - Trend
  - Seasonality or periodicity
  - Autocorrelation and temporal dependency
- Stochastic behavior of time series
  - Change points
  - Volatile behavior





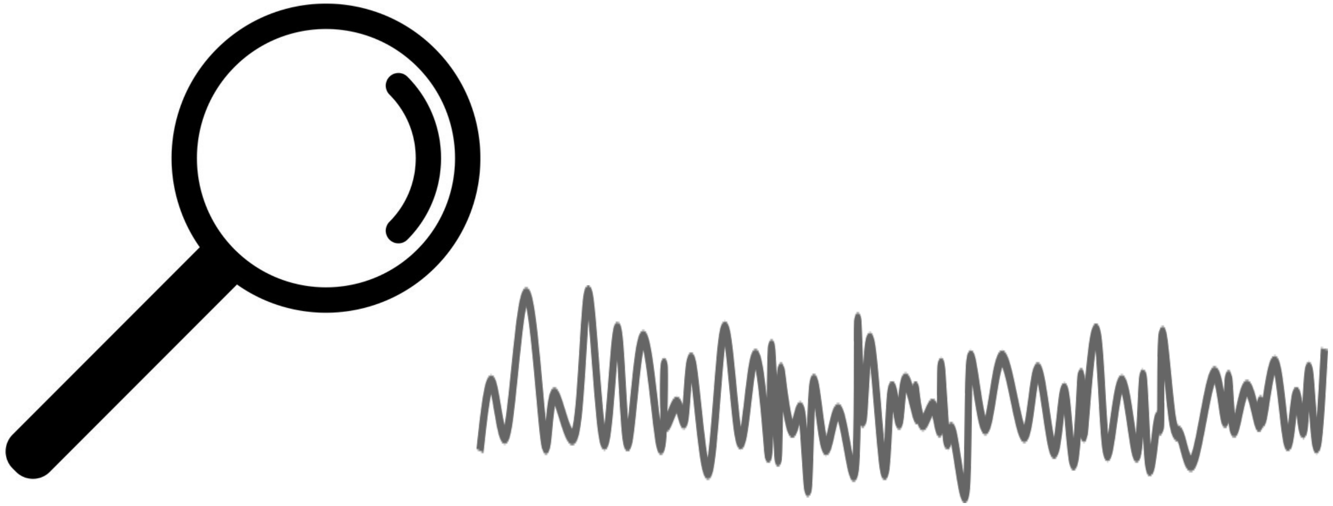


Objectives of time series analysis

- Exploratory data analysis
  - How does data behave over time? Trend, seasonality, stationarity
  - Understanding the nature of processes generating the data

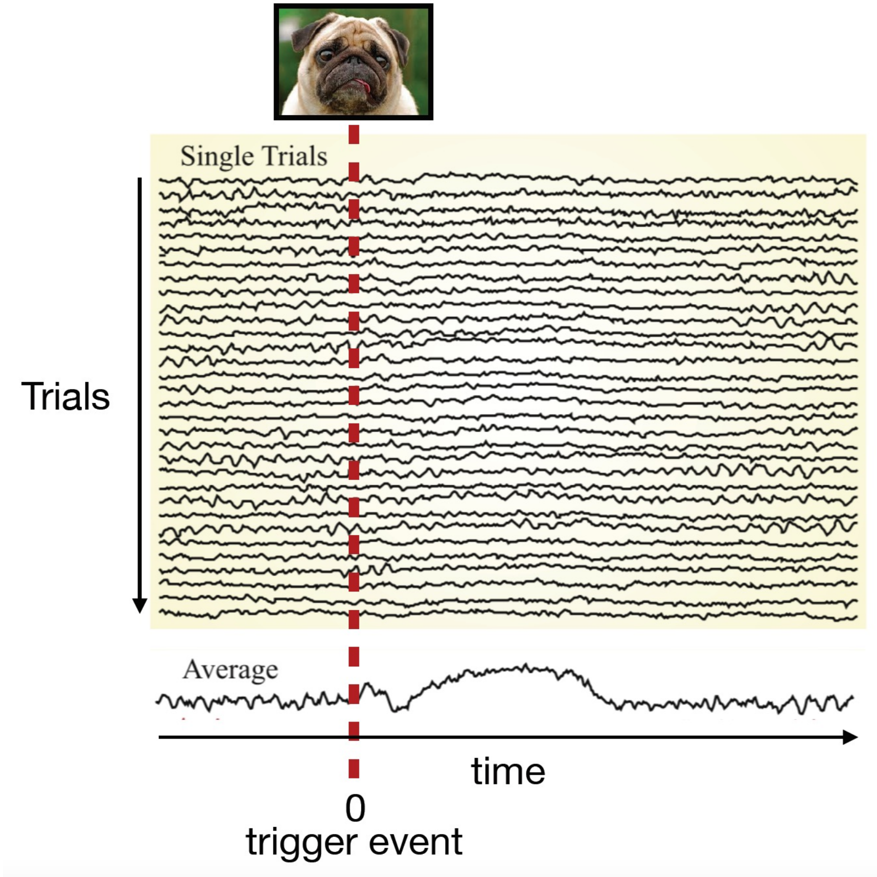
- Predictive data analysis
  - Estimate past, present, and future of data
  - Time series forecasting
  - Time series classification
  - Anomaly prediction

Explore the data!

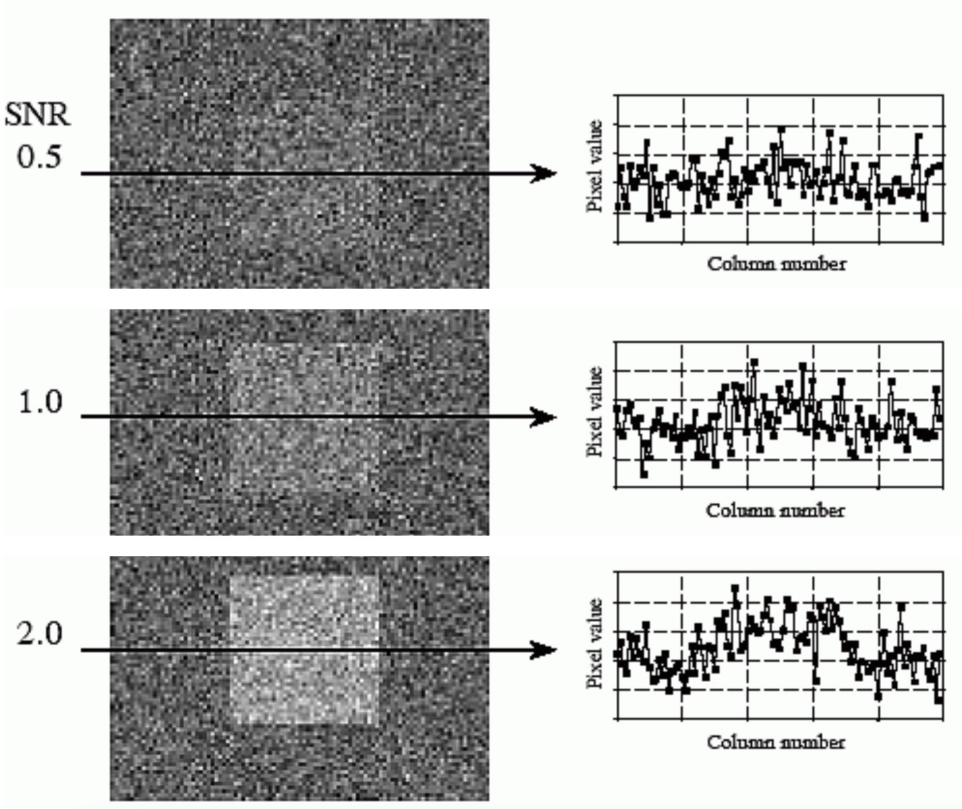


Signal and noise components

- Real world time-series are noisy.
- Signal-to-noise ratio (SNR) quantifies how noisy a signal is.



Constant noise with increasing signal



Root mean square (rms)

$$\text{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{n} x_i^2}$$

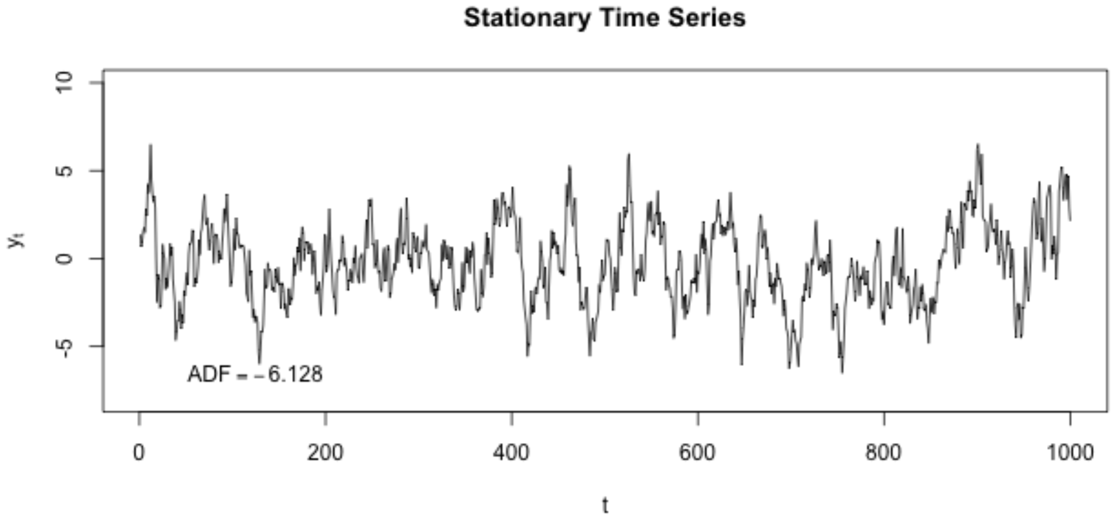
SNR in decibels (dB)

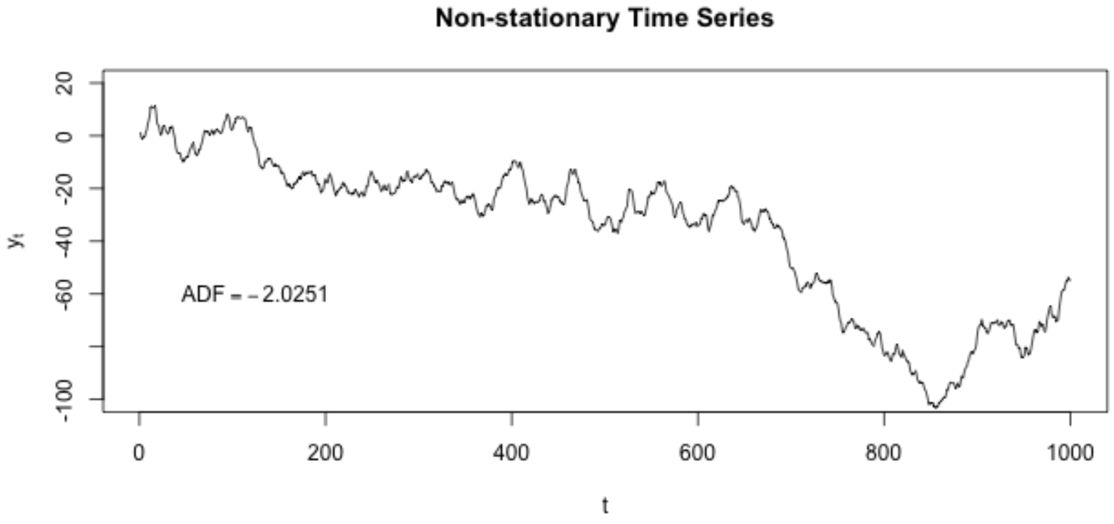
$$\text{SNR} = 20\log_{10}\frac{rms_{signal}}{rms_{noise}}$$

Neural Signal Processing course by Richard Gao: https://github.com/rdgao/COGS118C

Stationary vs non-stationary time series

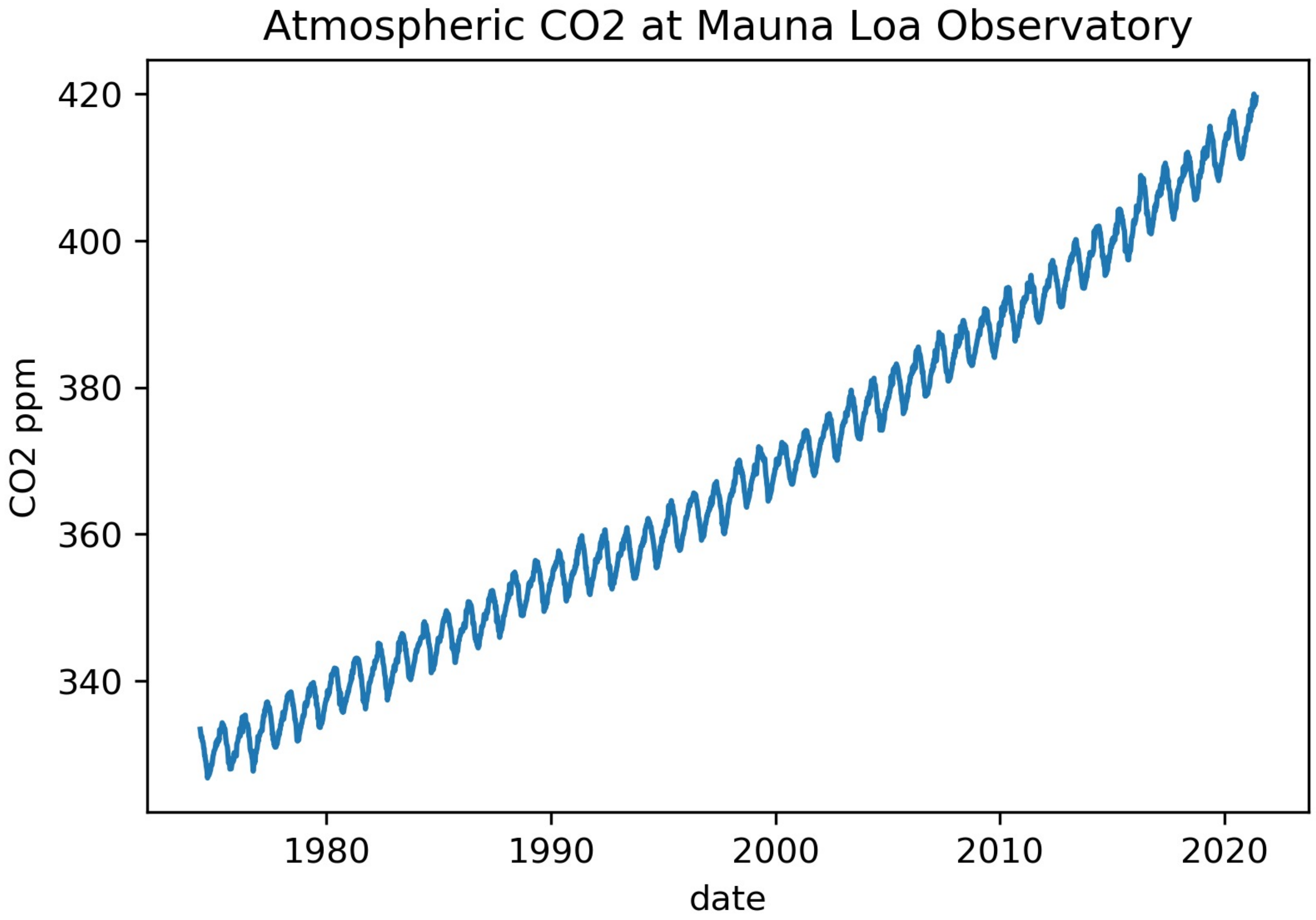
- Stationary time series
  - Mean, standard deviation, and structure of autocorrelation function are constant over time





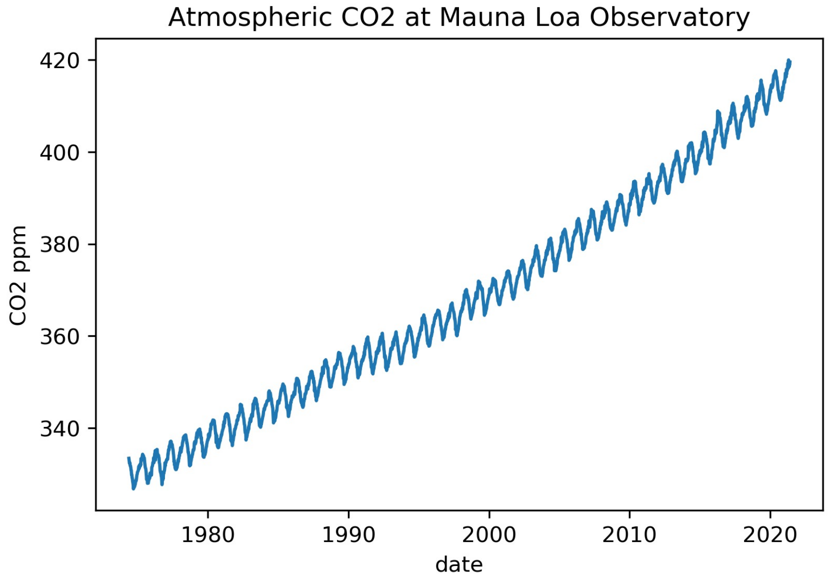
Example dataset

- Atmospheric CO_2 at Mauna Loa Observatory
- Trend: change in the mean

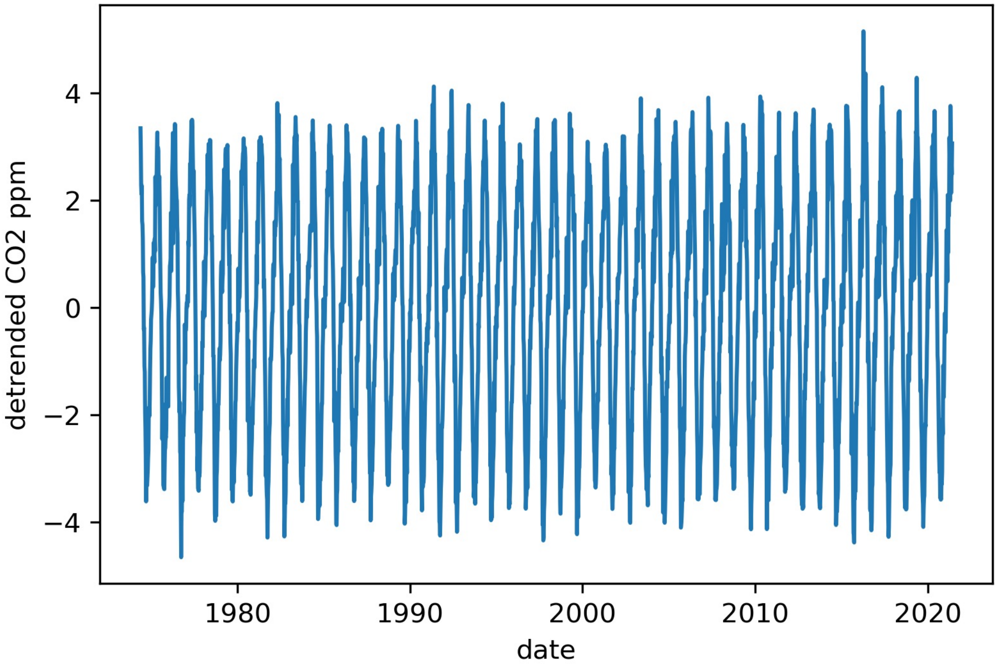


Example dataset

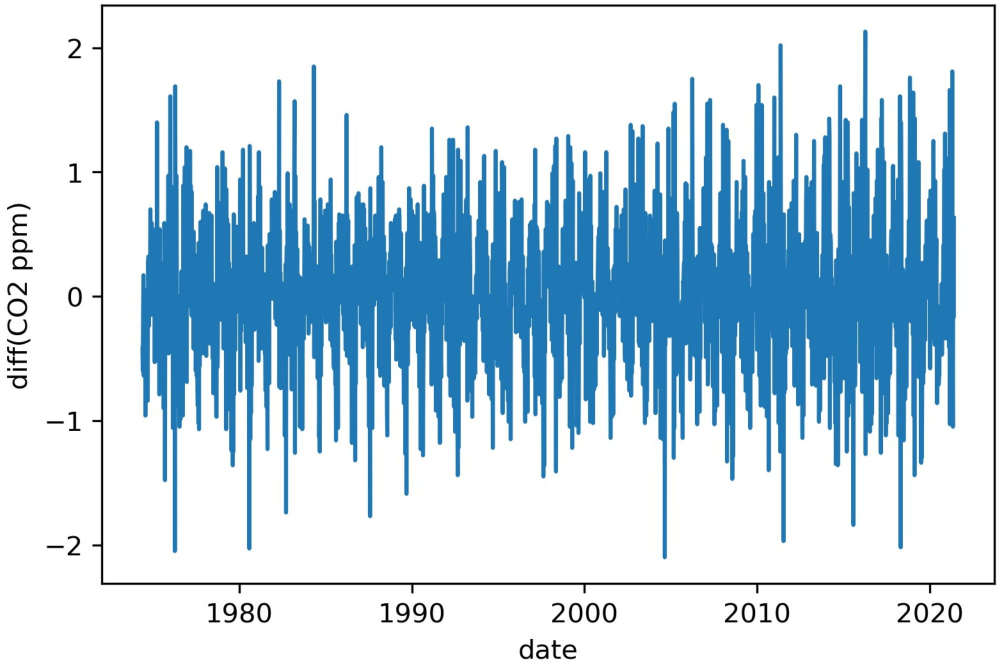
- Detrending
    - Remove the linear trend
    - Differencing



Removing linear trend



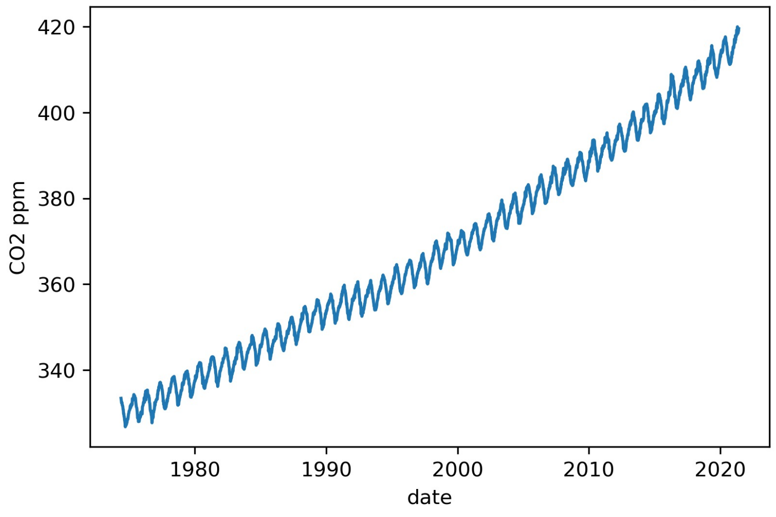
Differencing



Example dataset

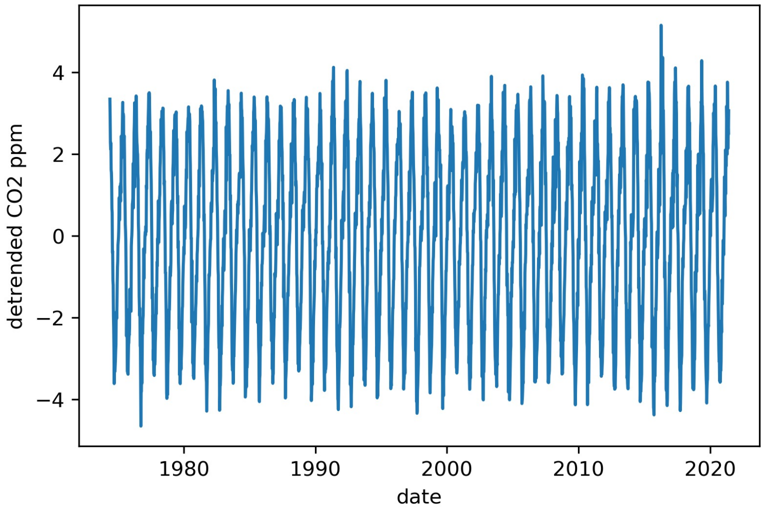
- Stationarity tests:
  - ADF: Augmented Dickey Fuller test
  - KPSS: Kwiatkowski-Phillips-Schmidt-Shin test

original



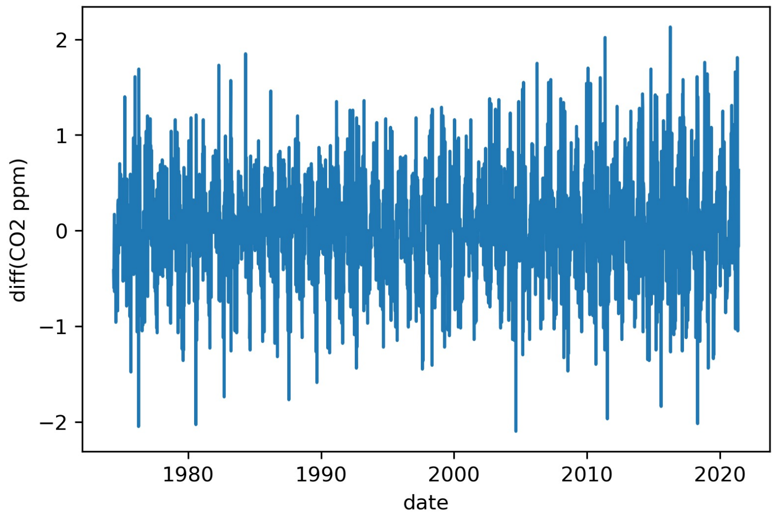
ADF Statistic: 0.46
p-value: 0.98

trend-stationary



ADF Statistic: -25.39
p-value: 0.0

difference-stationary

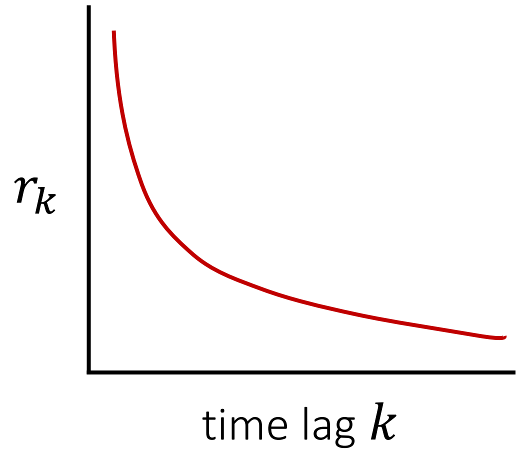


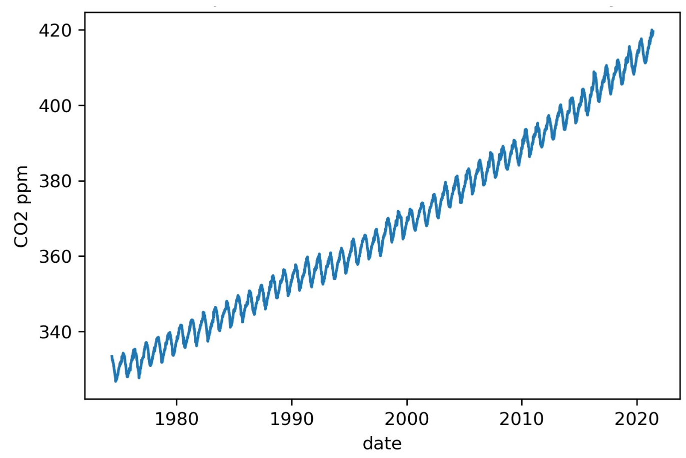
ADF Statistic: -15.92
p-value: 7.89e-29

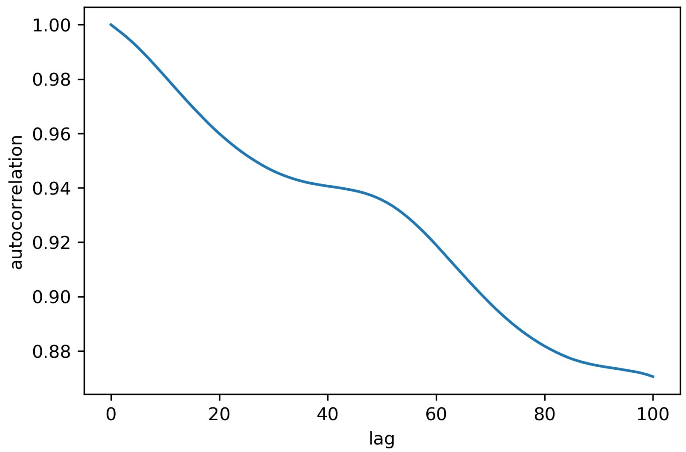
Example dataset

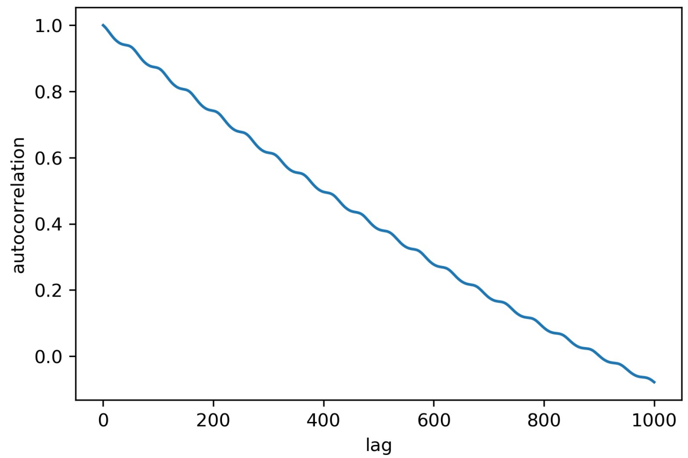
- Seasonality and periodicity
- Autocorrelation
  - Non-randomness in the data
  - Model selection for data fitting

Autocorrelation function
$$r_k = \frac{\sum_{i=1}^{N-k}(Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}$$



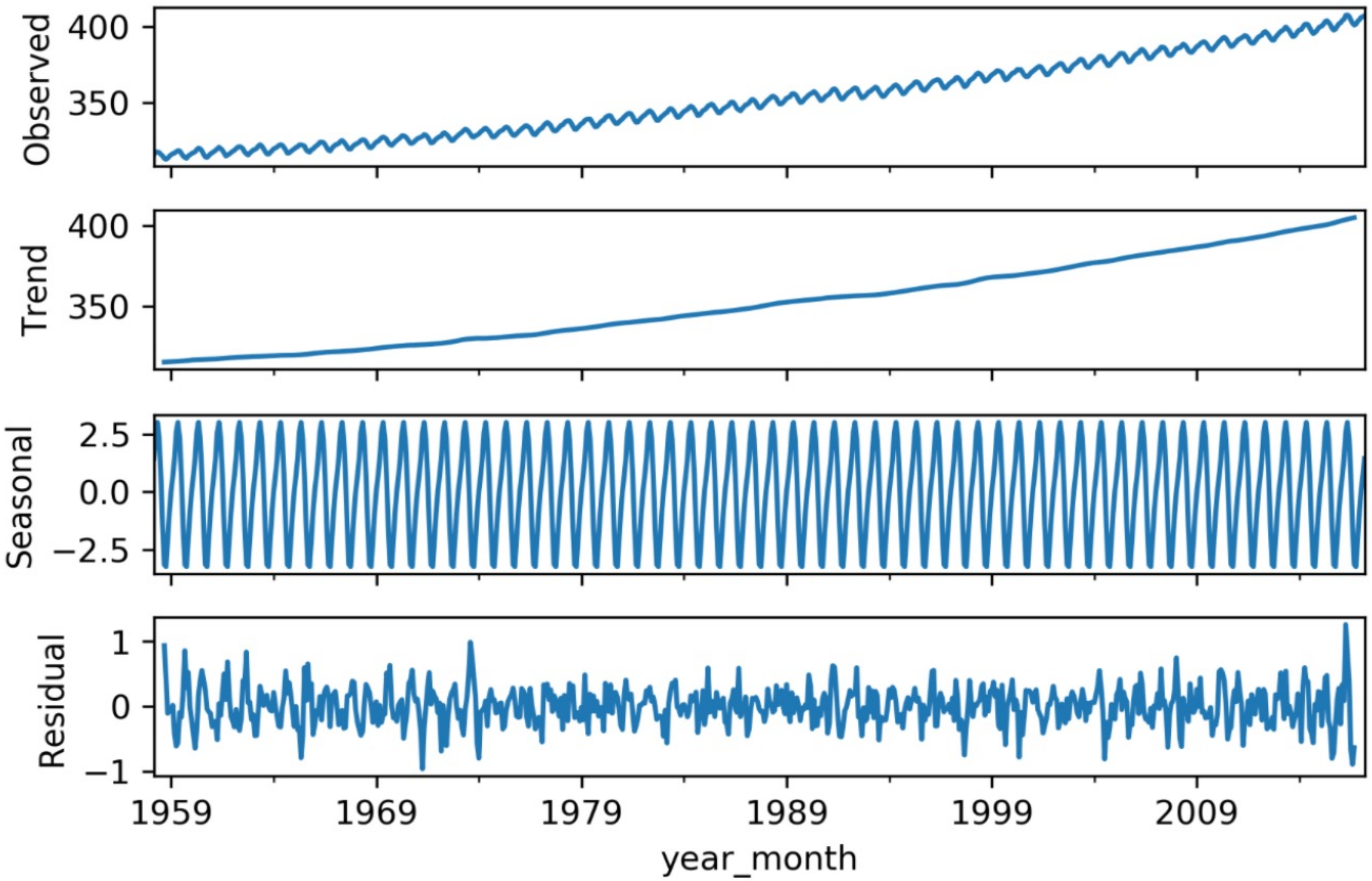






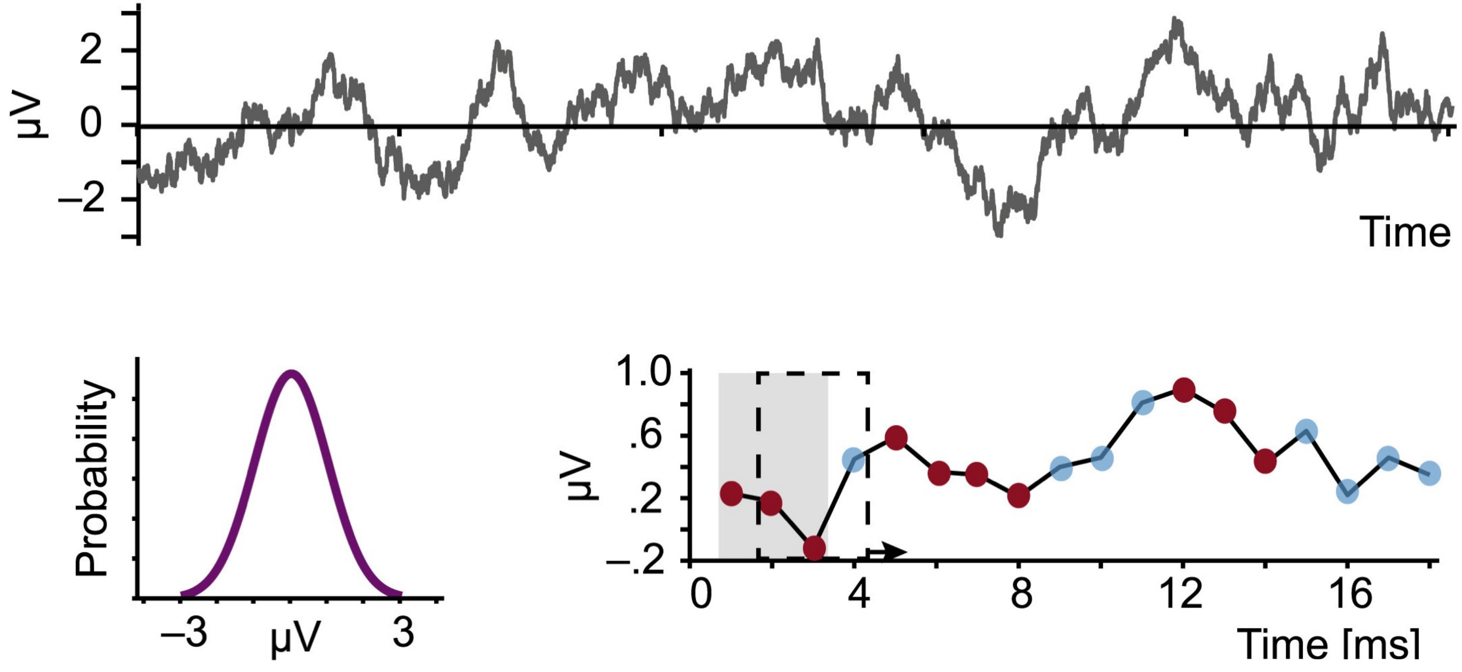
Example dataset

- Decompose signal into different components



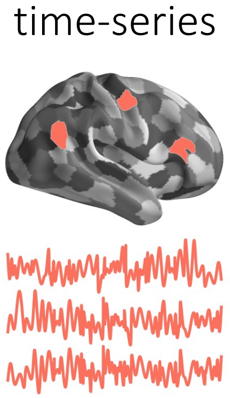
Measures of time series

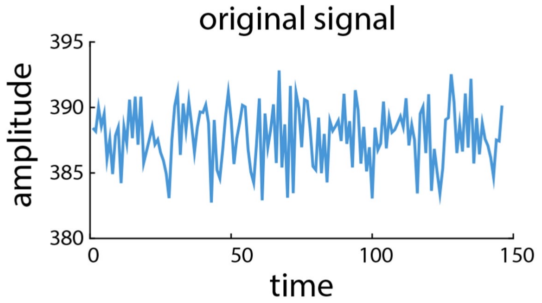
- Stationarity
- Distribution: mean, variance, skewness, kurtosis
- Temporal dependencies: autocorrelation, temporal entropy



Waschke et al., 2021, *Neuron*

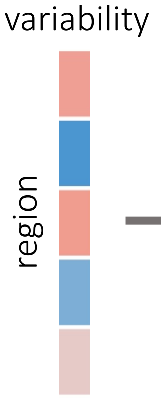
Measures of time series

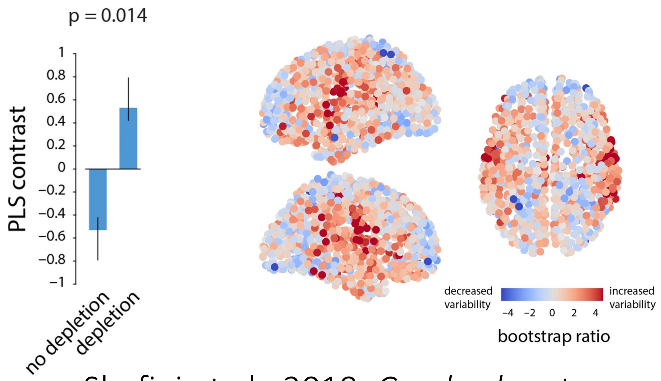




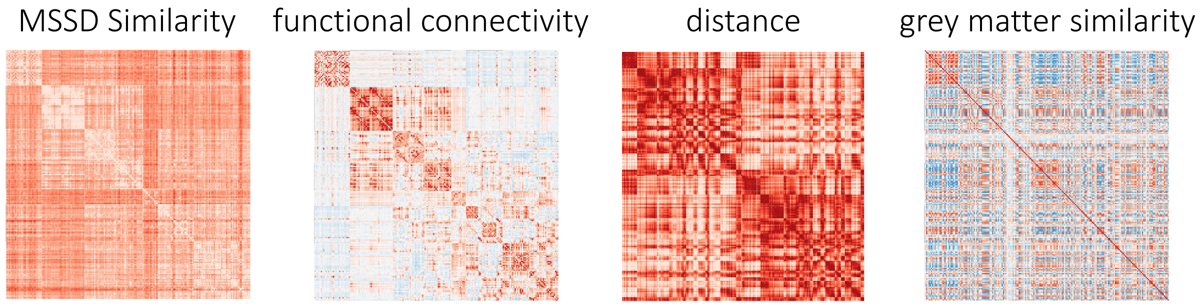
$$SE(m, r, N) = \ln \frac{\sum_{i=1}^{N-m} n_i^m}{\sum_{i=1}^{N-m} n_i^{m+1}}$$

$$MSSD = \sqrt{\frac{\sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}{n-1}}$$



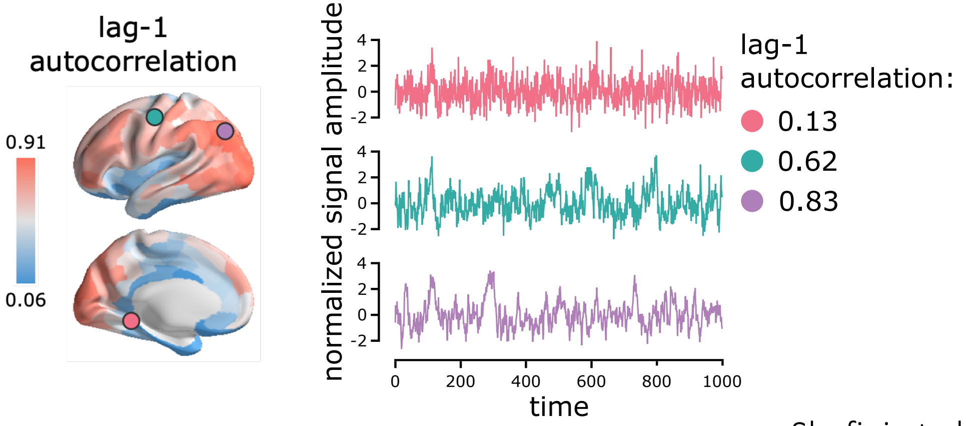


Shafiei et al., 2019, *Cerebral cortex*

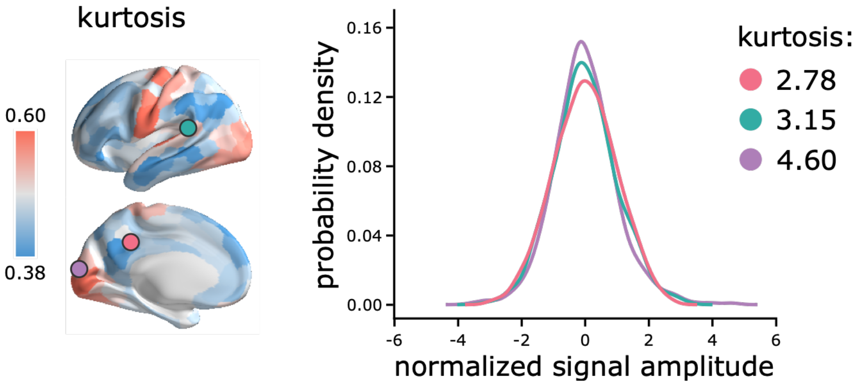


Baracchini et al., 2021, *NeuroImage*

Autocorrelation



Data distribution



Shafiei et al., 2020, *eLife*

Time series forecasting

- **Objective:** Predict time-series values for future time points from past values.
- AutoRegressive models: AR(p)

lagged time series

time series at time t $$y_t = B_0 + B_1 \underline{y_{t-1}} + B_2 \underline{y_{t-2}} + \cdots + B_p \underline{y_{t-p}} + e_t$$ forecast error term

coefficients

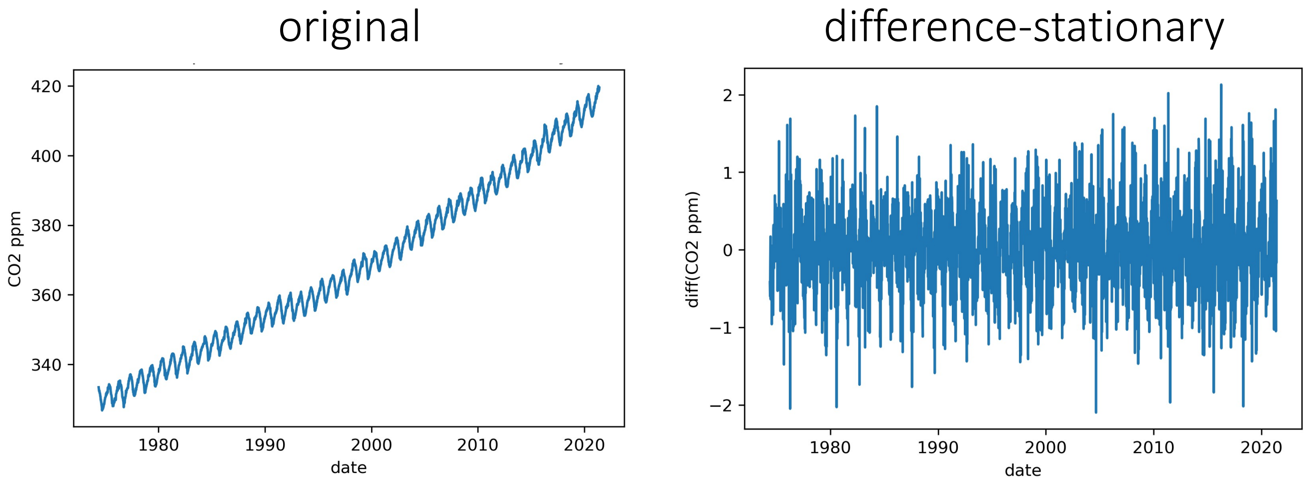
- Moving average models: MA(q)

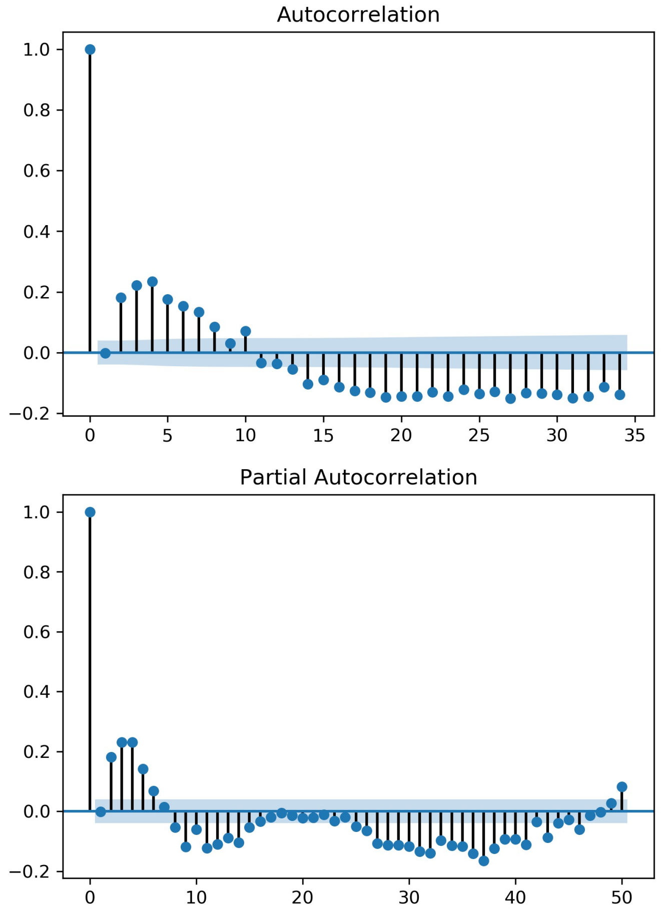
$$y_t = \mu + \underline{e_t} + C_1 \underline{e_{t-1}} + C_2 \underline{e_{t-2}} + \cdots + C_q \underline{e_{t-q}}$$

- Autoregressive Moving Average models: ARMA(p,q)
- Autoregressive Integrated Moving Average models: ARIMA(p,d,q)
- Seasonal ARIMA models: SARIMA
- ...

Time series forecasting

- Parameter selection:
  - MA model parameter: Autocorrelation function (ACF)
  - AR model parameter: Partial autocorrelation function (PACF)
  - Grid search

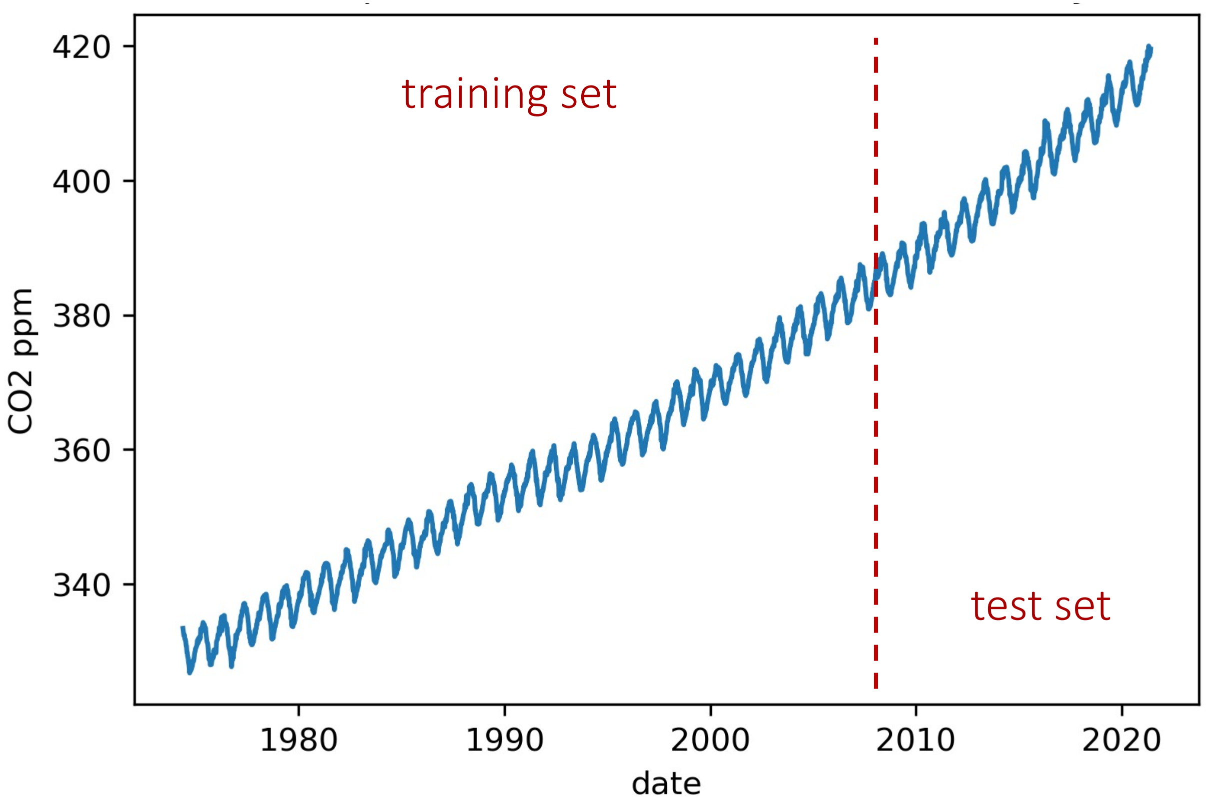




Time series forecasting

- Forecasting using AutoRegressive model from Python's *statsmodels*

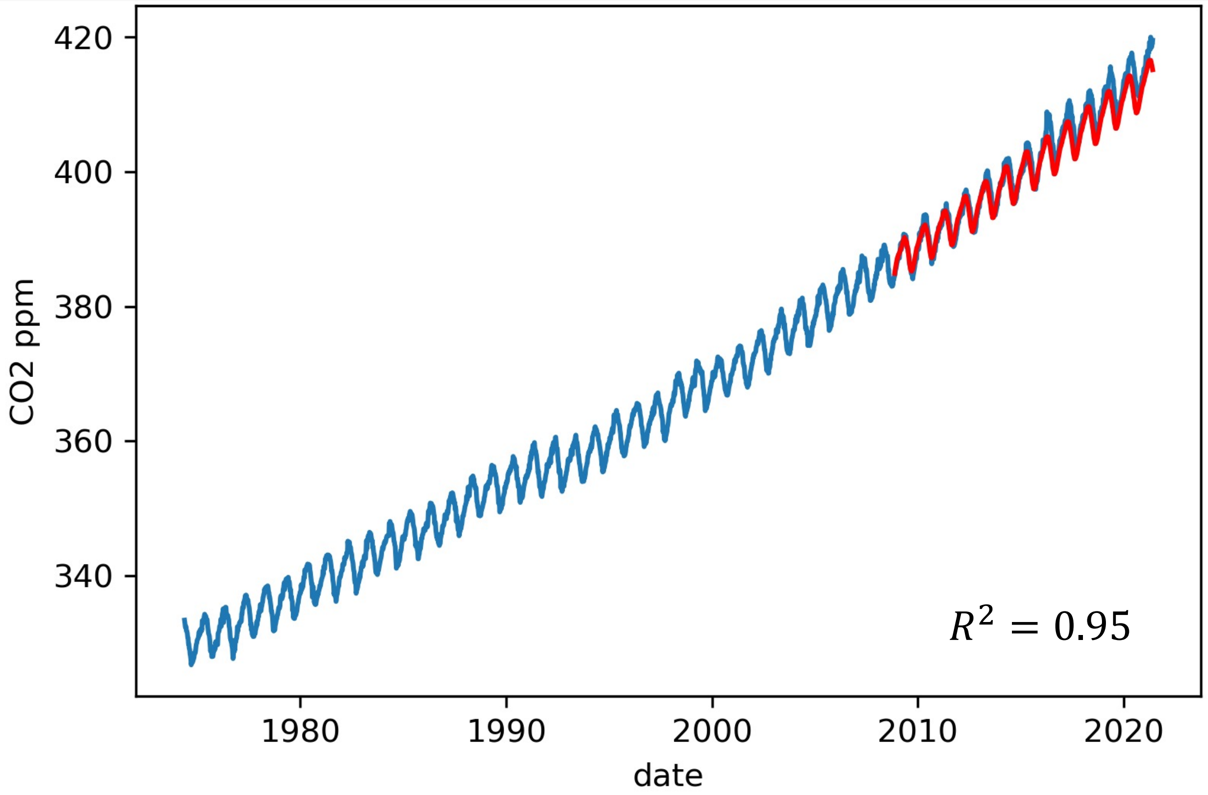
```
from statsmodels.tsa import ar_model
# train AutoRegressive model using training set
ar_model = ar_model.AR(train)
ar_model_fit = ar_model.fit(maxlag=50)
```



Time series forecasting

- Forecasting using AutoRegressive model from Python's *statsmodels*

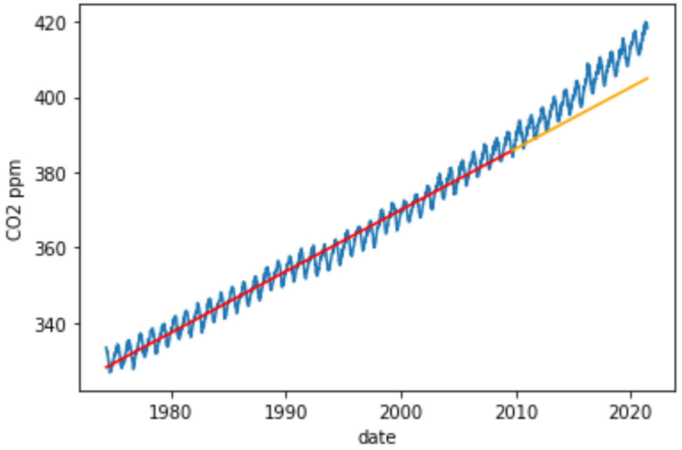
```
# predict and plot time-series values for test data using test set index
predicted_values_ar = ar_model_fit.predict(test.index[0], test.index[-1])
plt.plot(ppm)
plt.plot(predicted_values_ar, 'red')
plt.xlabel('date')
plt.ylabel('CO2 ppm')
```



Time series forecasting

- Forecasting using regression analysis and AR model

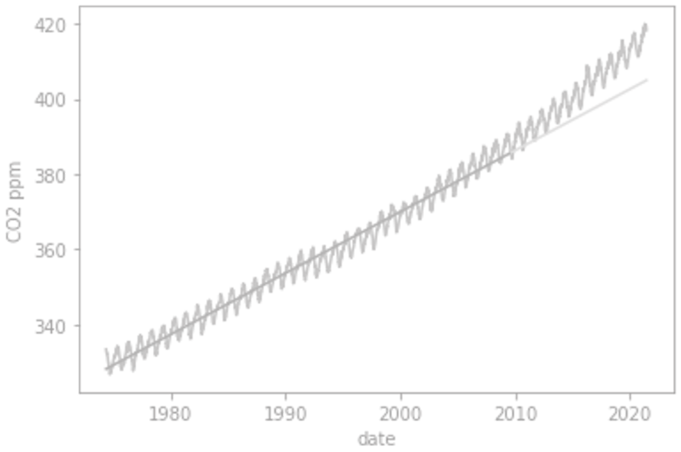
linear trend



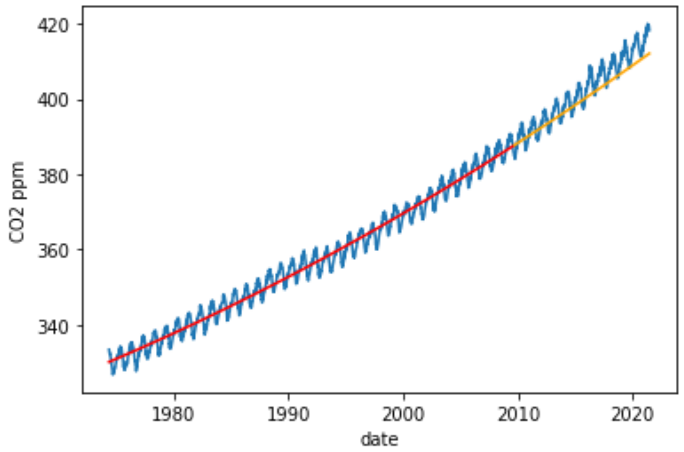
Time series forecasting

- Forecasting using regression analysis and AR model

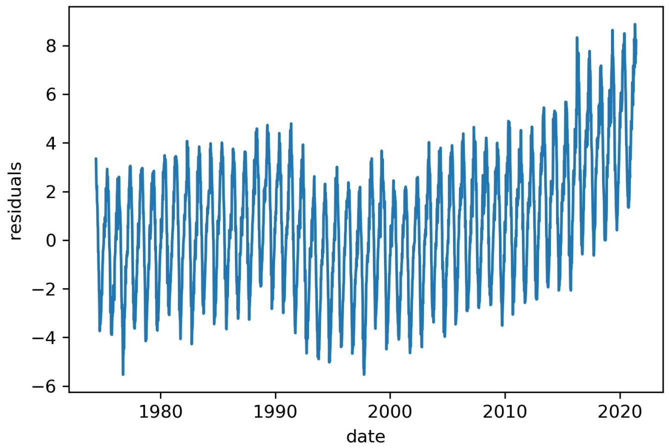
linear trend



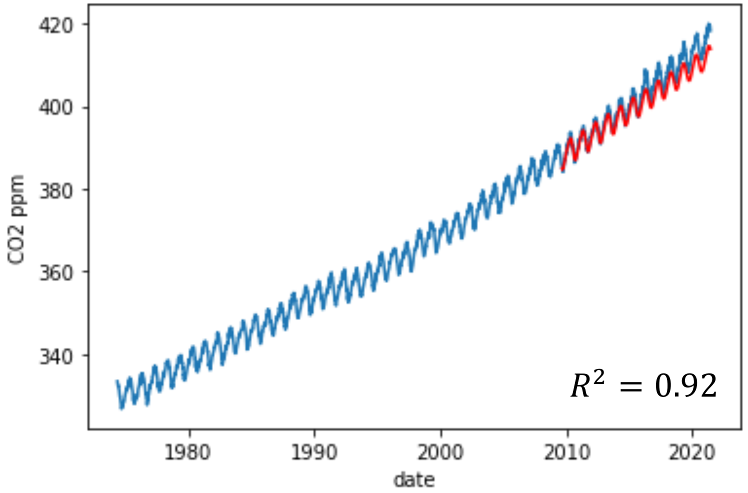
polynomial trend



detrending with polynomial trend

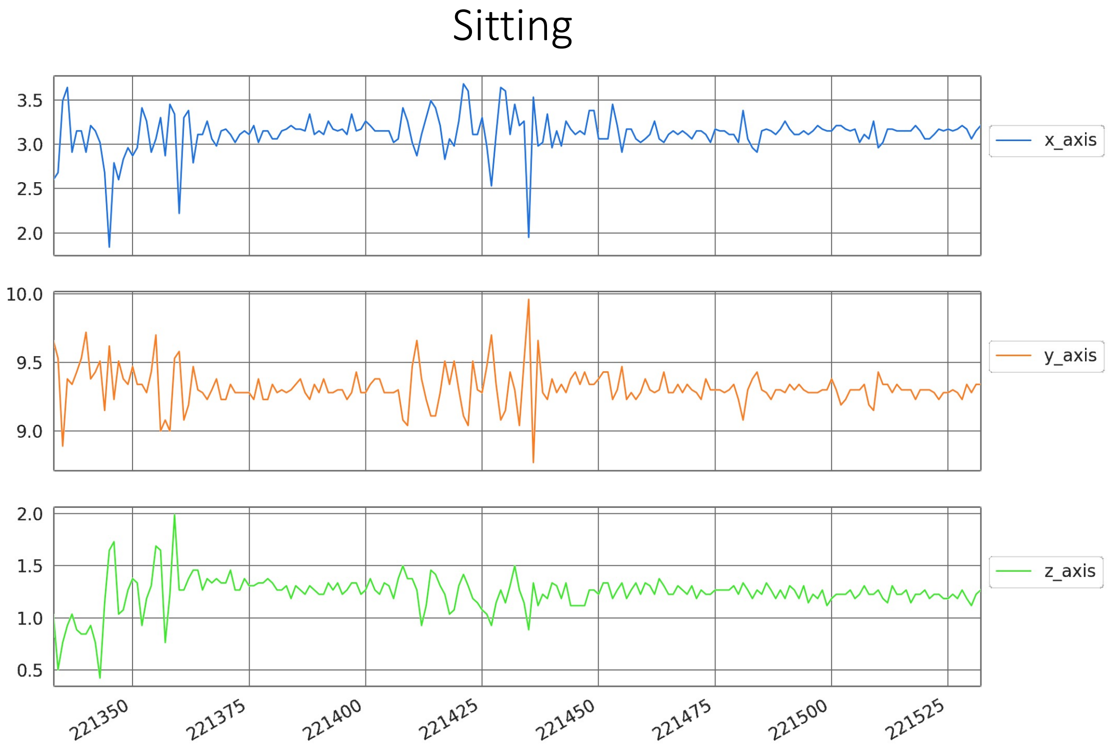


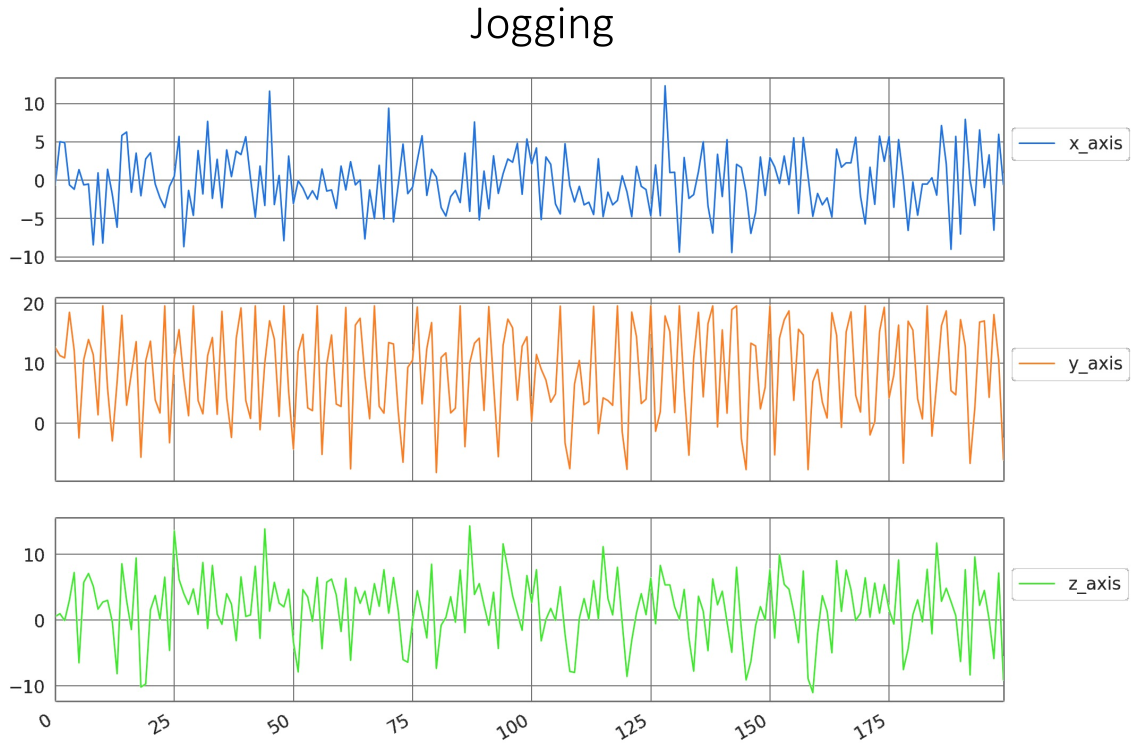
AR model on detrended data



Time series classification

- Classify time series into different categories

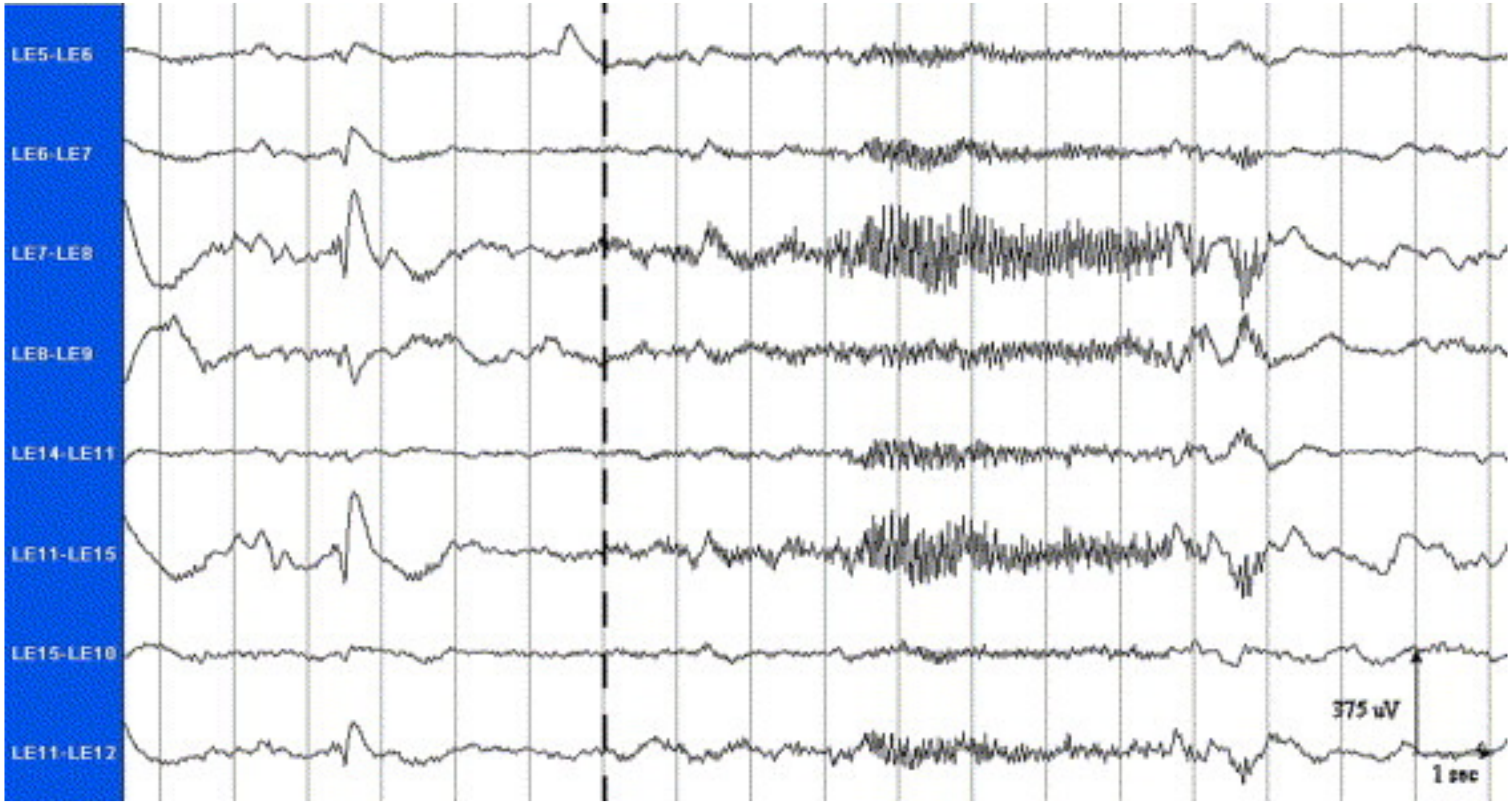




Time series classification

- Classify time series into different categories

Epileptic EEG recording

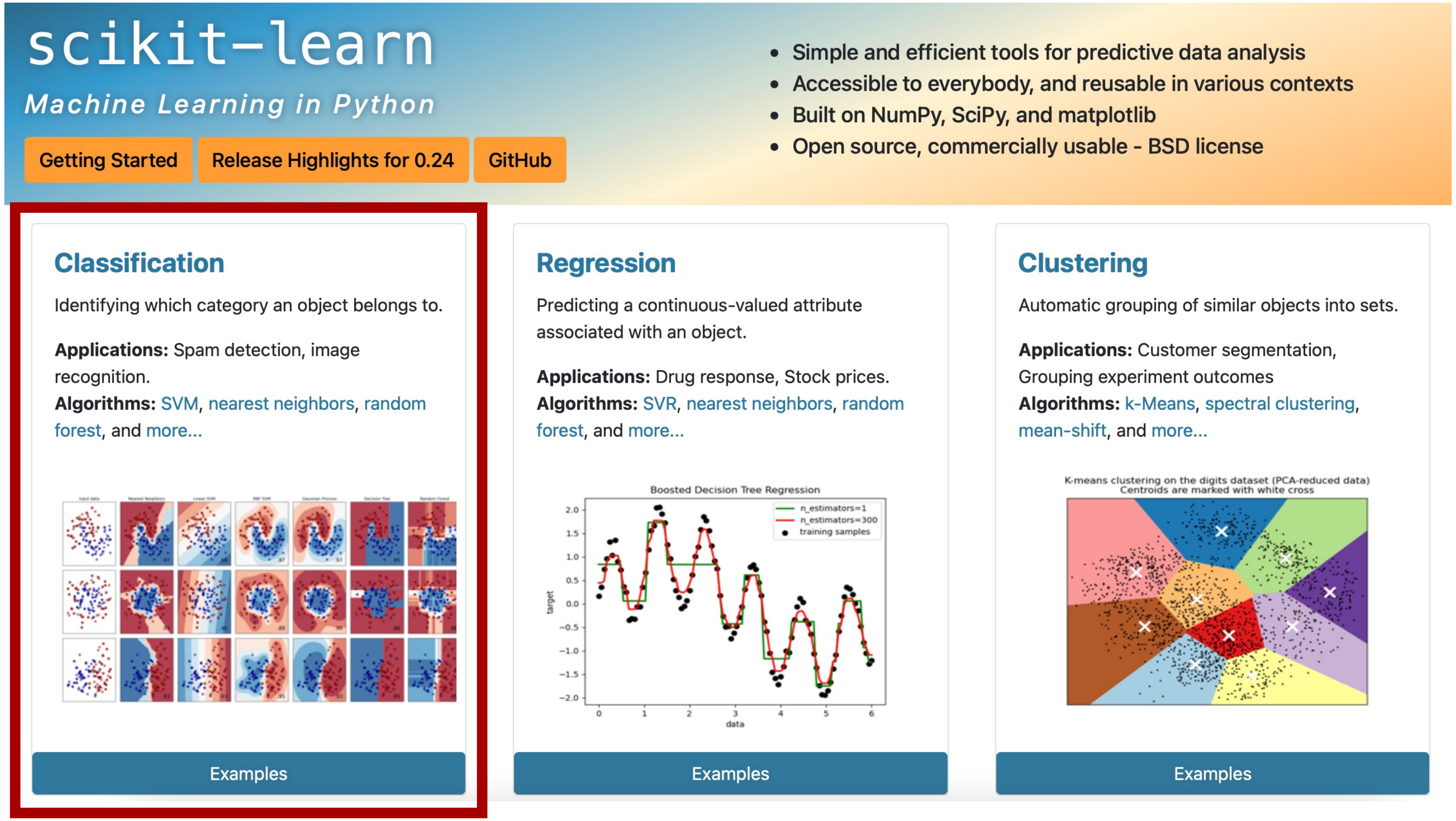


Grewal & Gotman, 2005, *Clin. Neurophysiol.*

Time series classification

- Classify time series into different categories
- Scikit-learn classification algorithms

https://scikit-learn.org/



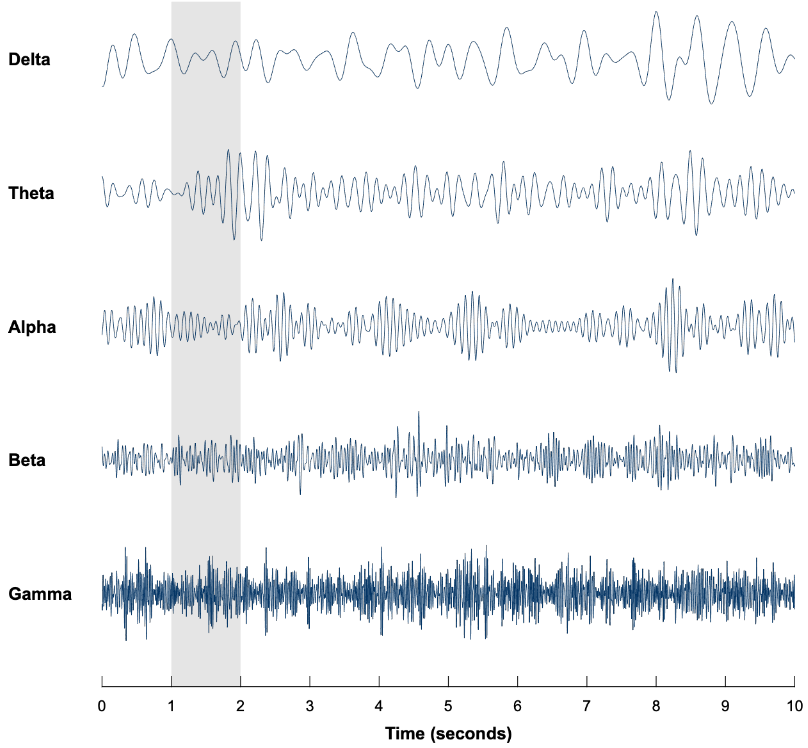
Spectral analysis

- Fourier transform

WIKIPEDIA
The Free Encyclopedia

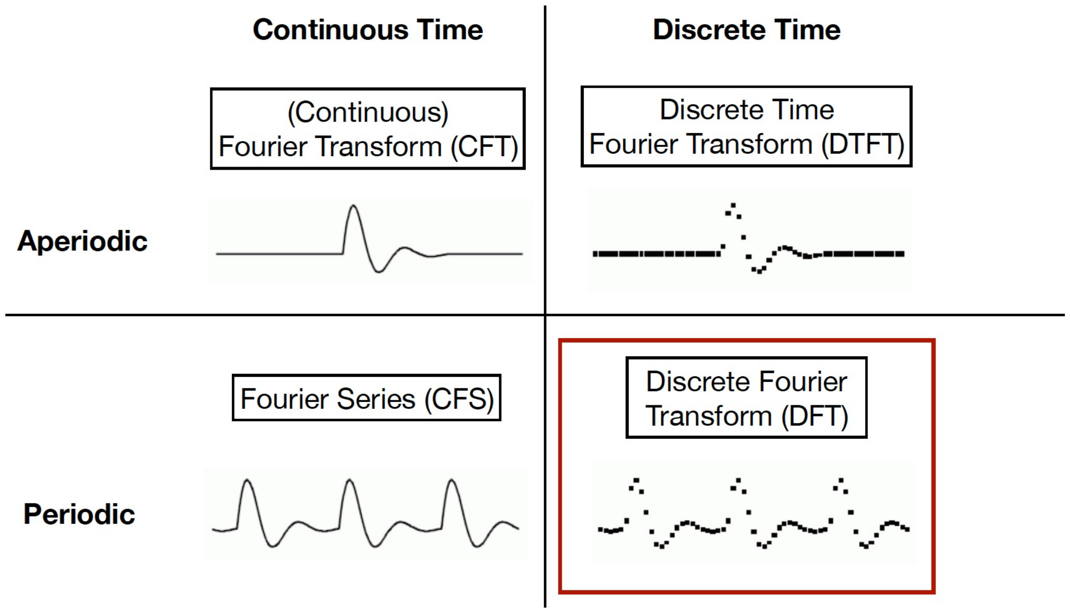
“Fourier transform is a mathematical transform that decomposes functions depending on space or time into functions depending on spatial or temporal frequency.”

Neural oscillation from EEG



https://en.wikipedia.org/wiki/Neural_oscillation

Family of Fourier Transforms



Neural Signal Processing course by Richard Gao
https://github.com/rdgao/COGS118C

Spectral analysis

- Discrete Fourier transform (DFT) transforms a sequence of N complex numbers $\{x_n\}$ to another sequence of complex numbers $\{X_k\}$:

Seq1: $\{x_n\} := x_0, x_1, \dots, x_{N-1}$ *n* is time

Seq2: $\{X_k\} := X_0, X_1, \dots, X_{N-1}$ *k* is frequency

DFT

$$X_k = \sum_{n=0}^{N-1} x_n . e^{-\frac{i2\pi}{N}kn}$$

$$= \sum_{n=0}^{N-1} x_n . \left[\cos\left(\frac{2\pi}{N}kn\right) - i.\sin\left(\frac{2\pi}{N}kn\right)\right]$$

Euler's formula
$e^{ix} = \cos(x) + i\,sin(x)$

- ➢ Both k and n range between ***0*** and ***N-1***.
- ➢ As many frequency points as time points!

- Parseval's theorem: Sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its Fourier Transform. For DFT:

$$\text{Power } P = \sum_{n=0}^{N-1} |x_n|^2 = \frac{1}{N}\sum_{k=0}^{N-1} |X_k|^2$$

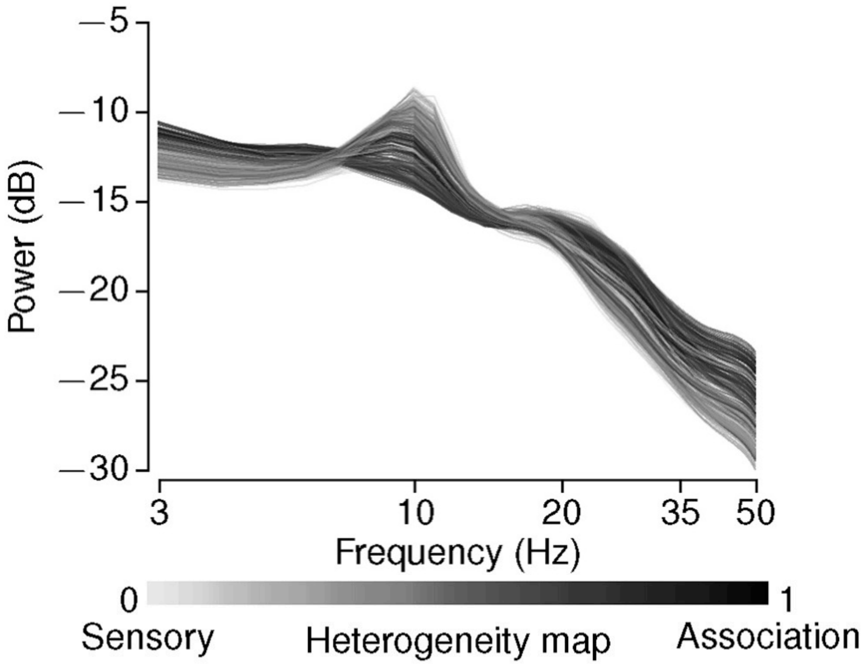
- ➢ If x_t measures voltage of a source in time t, power of x_t will be $(x_t)^2$.
- ➢ Thus, power is conserved in time and frequency domain!

- We can study power of a signal in frequency domain using Fourier Transform! (which is Power spectral analysis!)

Power spectral density

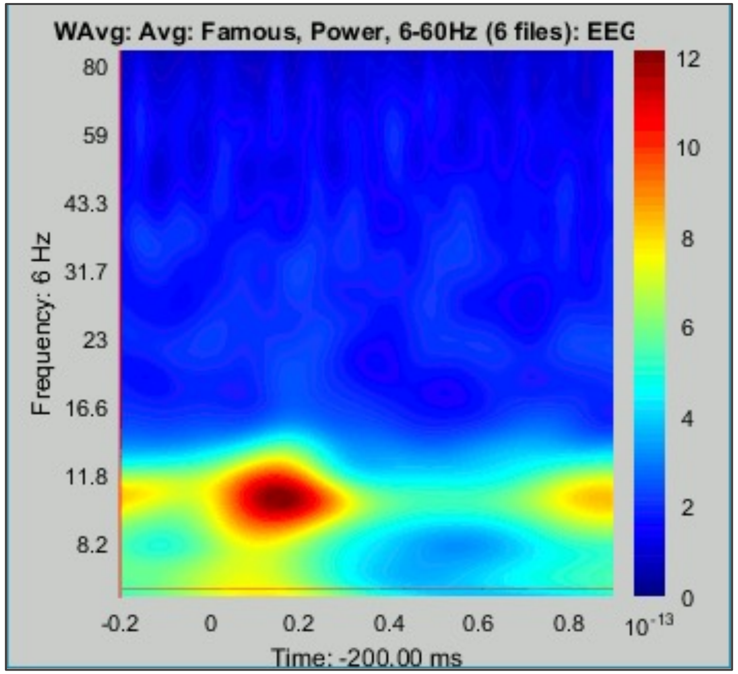
- There are different methods to measure power spectral density of a signal:
  - Welch's method: segments of time series with overlapping windows
  - Bartlet's method: segments of time series with non-overlapping windows

MEG power spectral density



Demirtas et al., 2019, *Neuron*

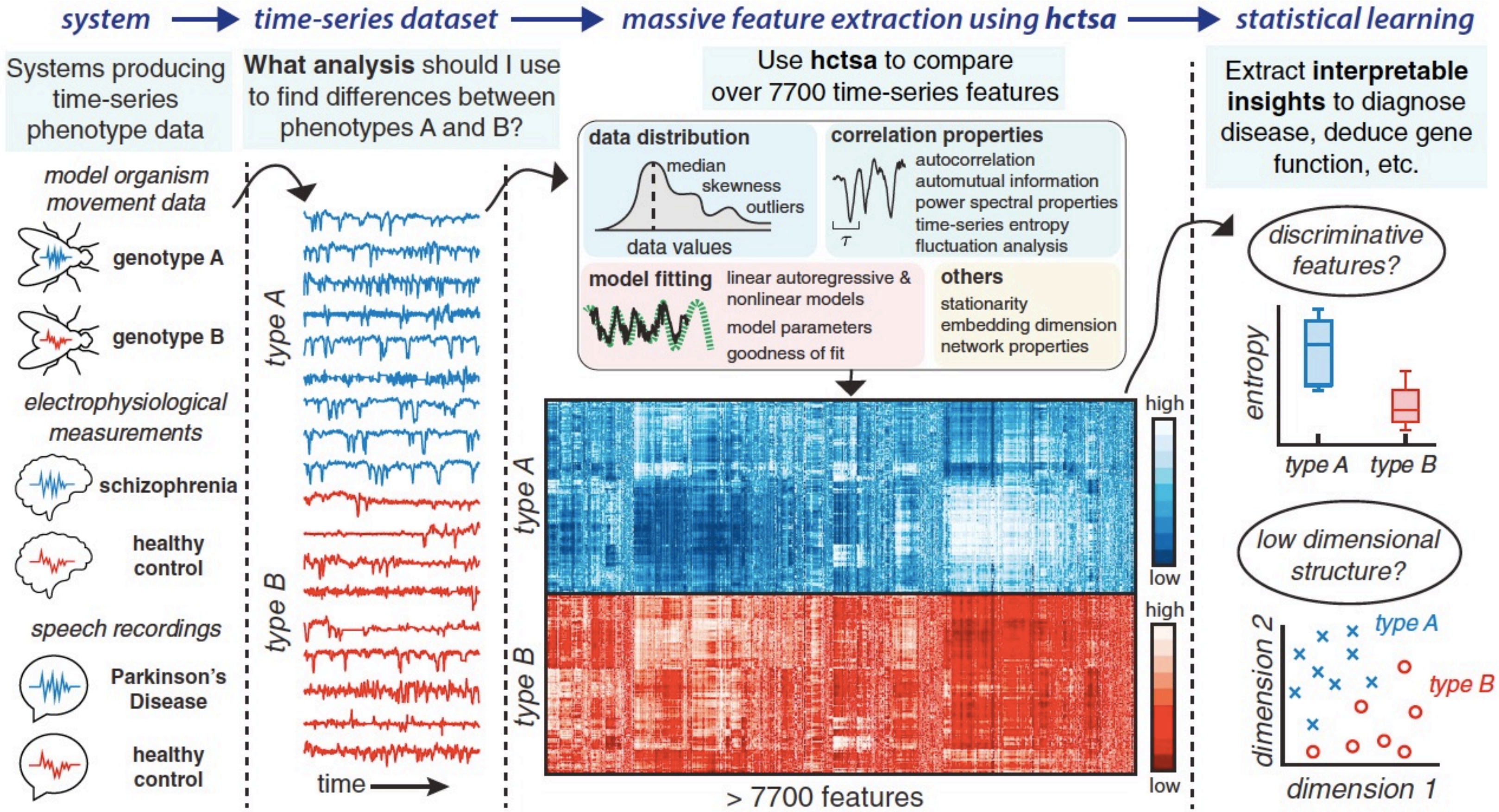
EEG time-frequency map



Brainstorm software package
https://neuroimage.usc.edu/brainstorm/

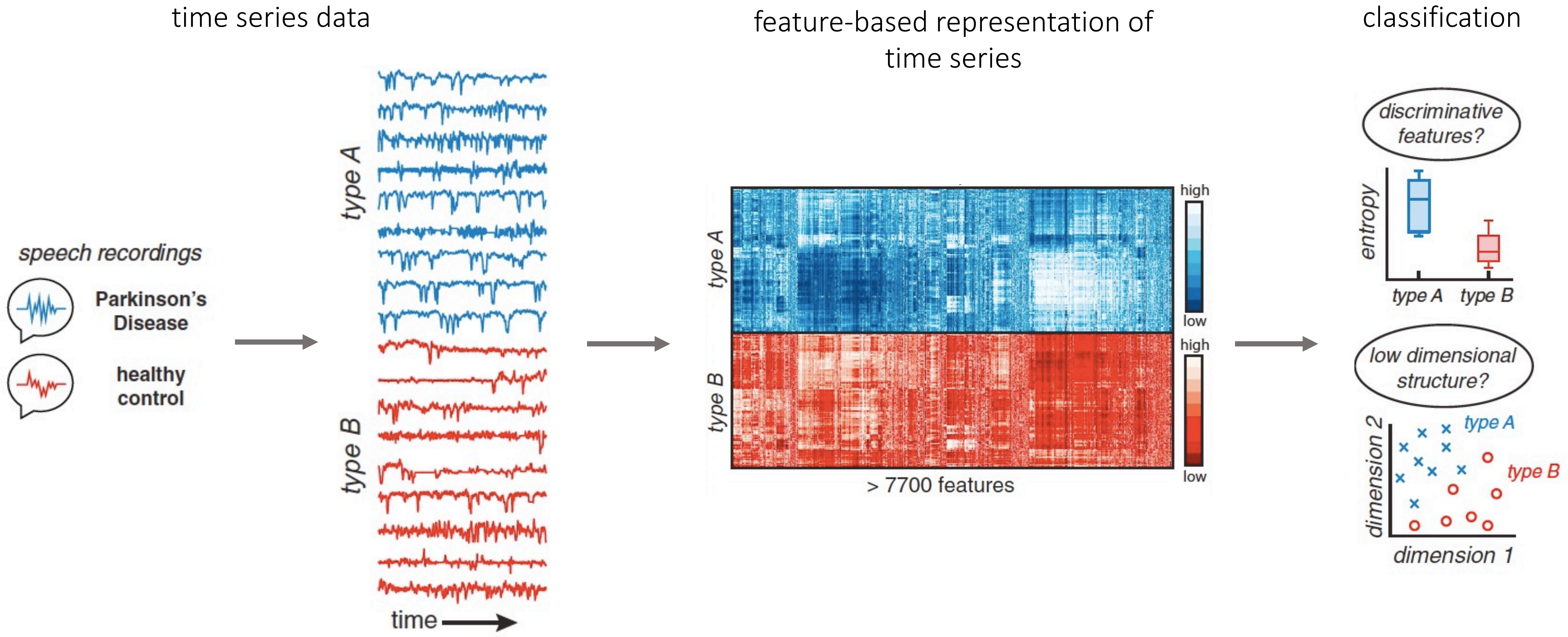
Time series phenotyping

- Feature-based representation of time series
- Extract several features from a time series
- Highly comparative time series analysis toolbox (hctsa)



Fulcher & Jones, 2017, *Cell Systems*

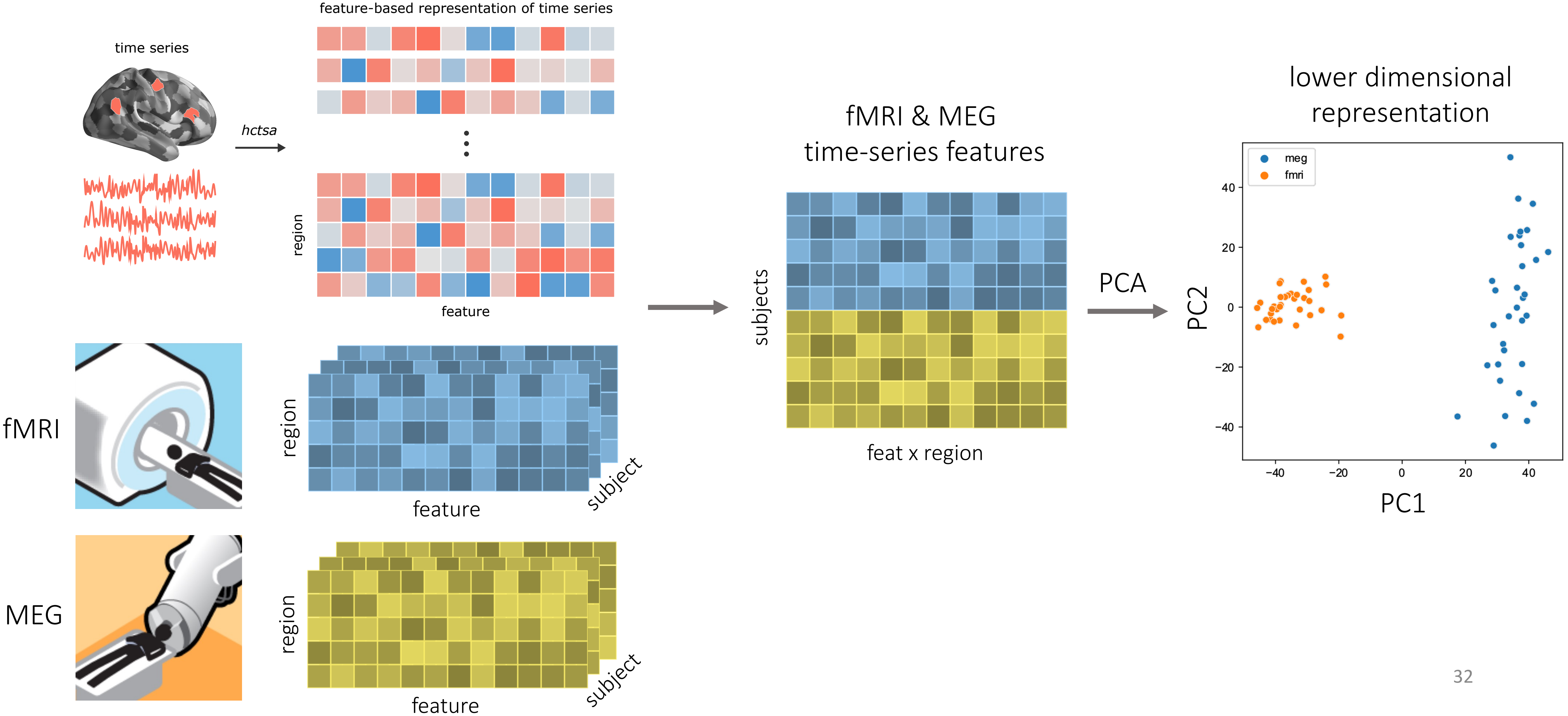
Time series classification with features



Fulcher & Jones, 2017, *Cell Systems*

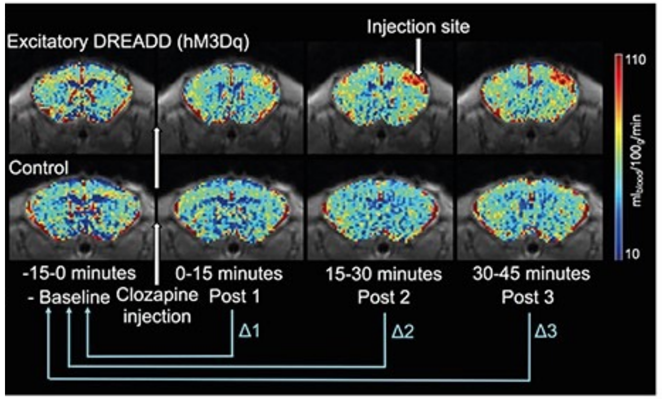
Time series classification with features

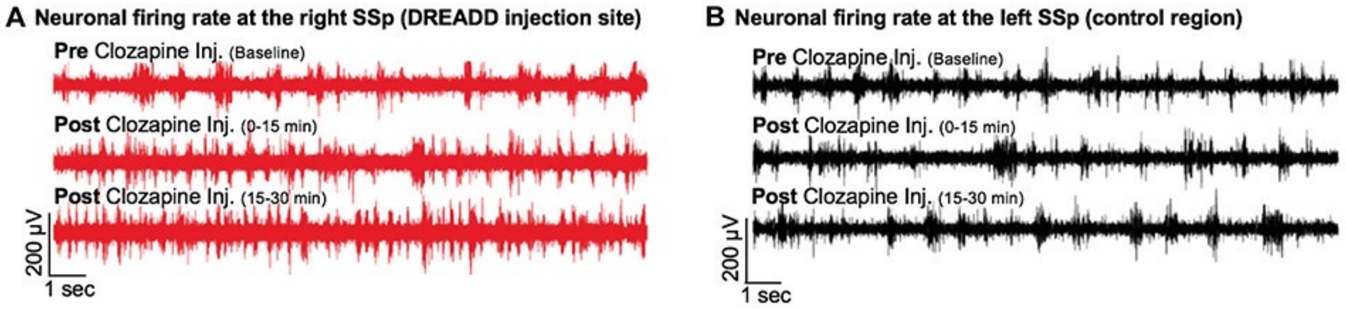
- Feature-based representation of time series from two imaging modalities



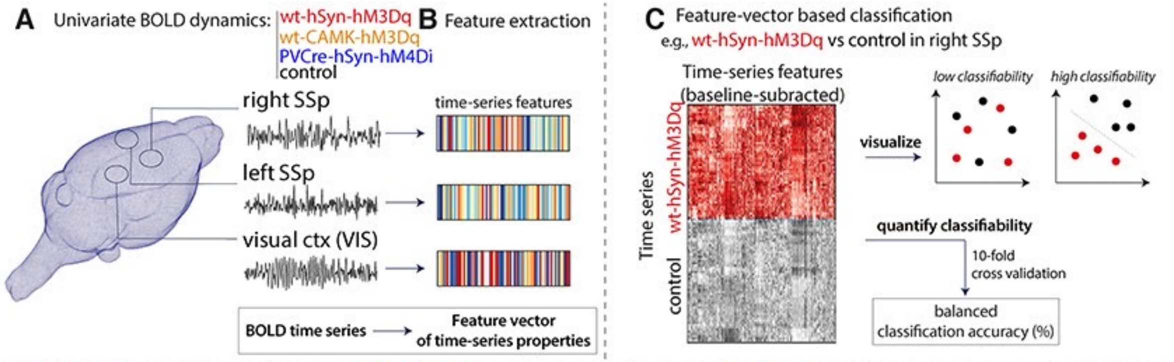
Time series classification with features

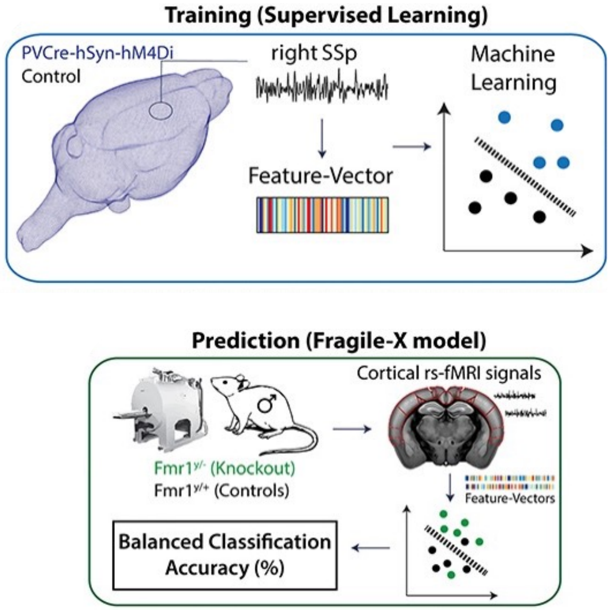
- Cortical E:I imbalance with designer receptor exclusively activated by designer drug (DREADD) injection

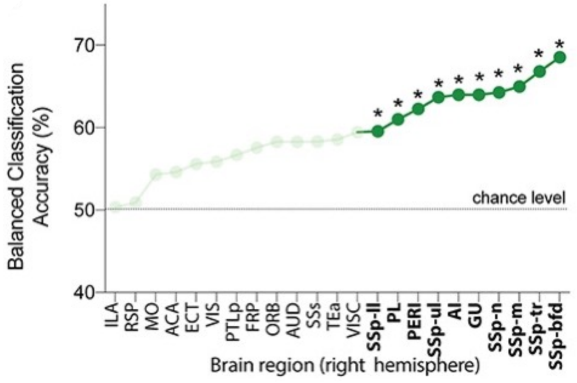




feature-based time-series representation







Markicevic et al., 2020, *Cerebral*