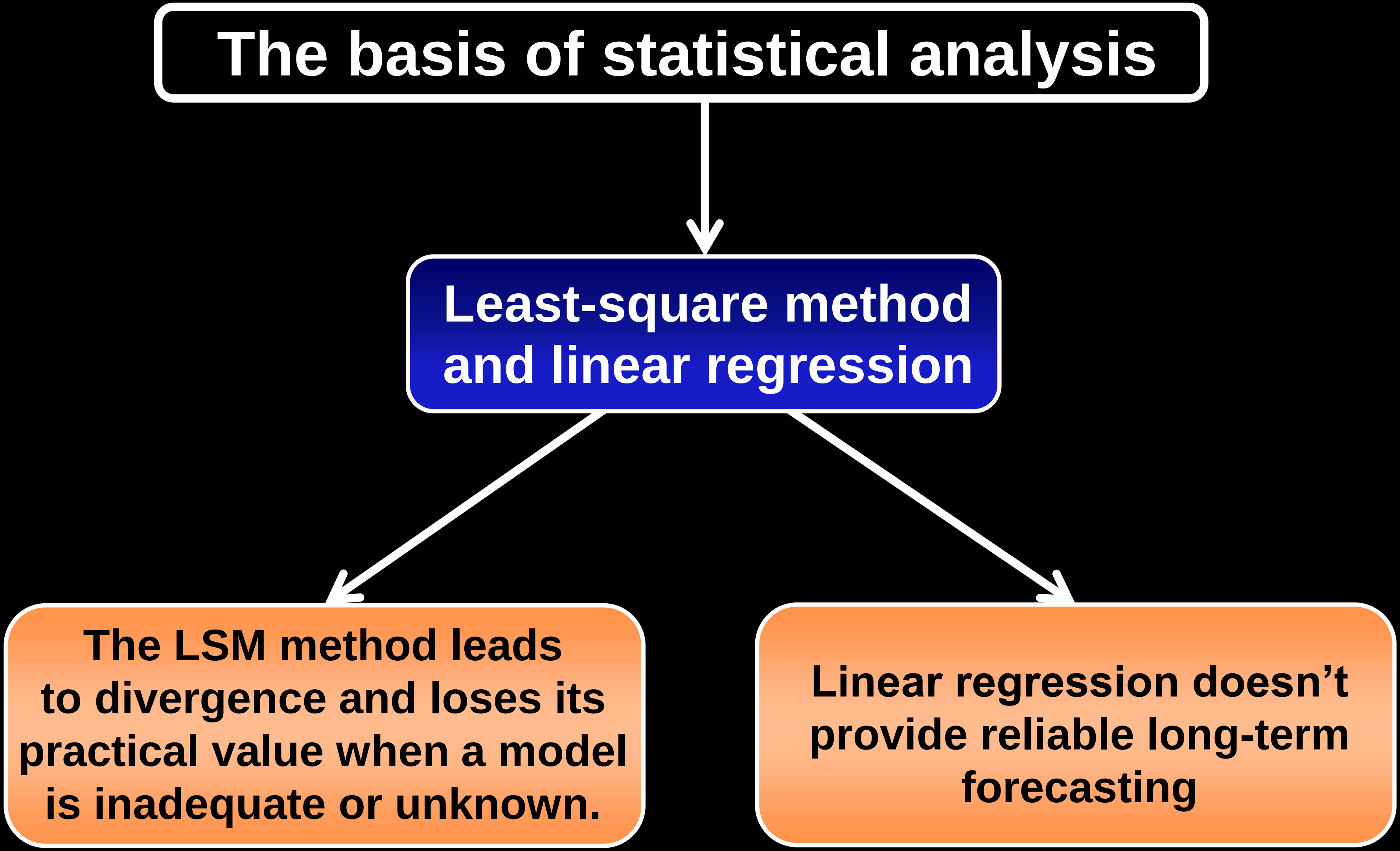
Skoltech

Skolkovo Institute of Science and Technology

“Experimental Data Processing”

Topic 2 "Quasi-optimal approximation under uncertainty”

Tatiana Podladchikova
Term 1B, October 2017
t.podladchikova@skoltech.ru



The basis of statistical analysis

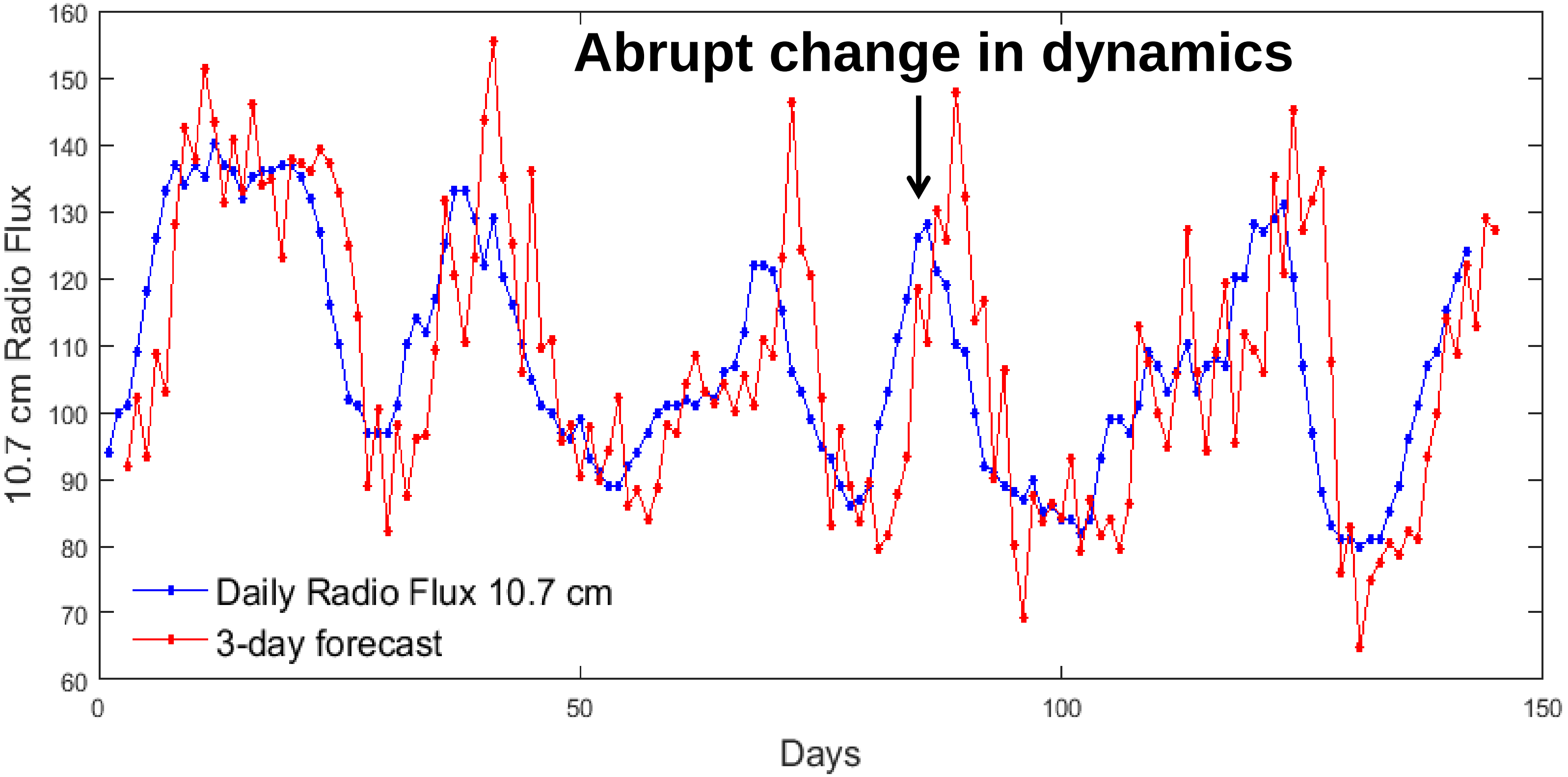
Least-square method and linear regression

The LSM method leads to divergence and loses its practical value when a model is inadequate or unknown.

Linear regression doesn’t provide reliable long-term forecasting

Linear regression doesn’t provide long-term forecasting

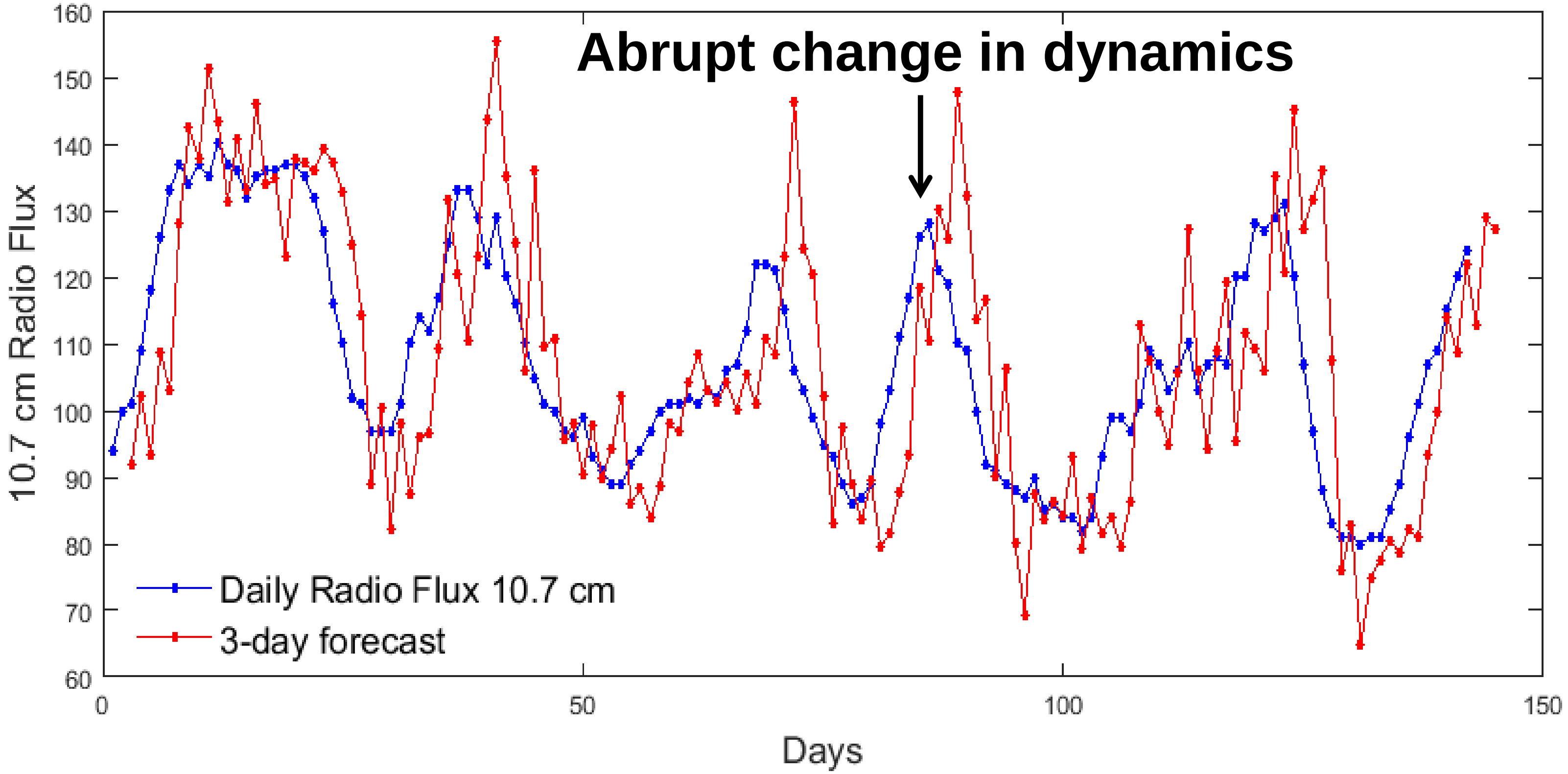
3-day 10.7 cm radio flux forecast based on a linear regression



Changes in dynamics of a process leads to great increase of forecasting errors

Linear regression doesn’t provide long-term forecasting

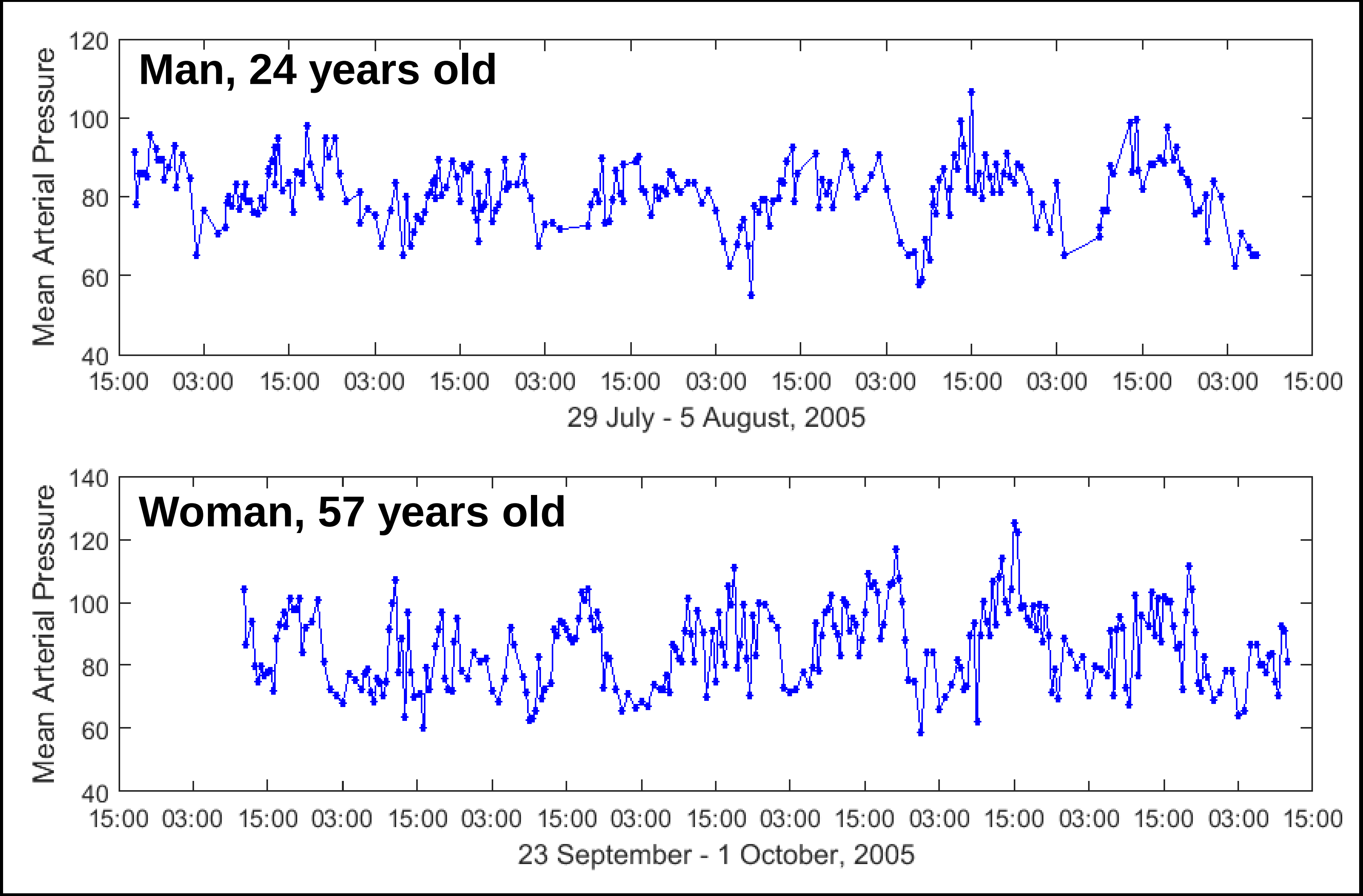
3-day 10.7 cm radio flux forecast based on a linear regression



Changes in dynamics of a process leads to great increase of forecasting errors

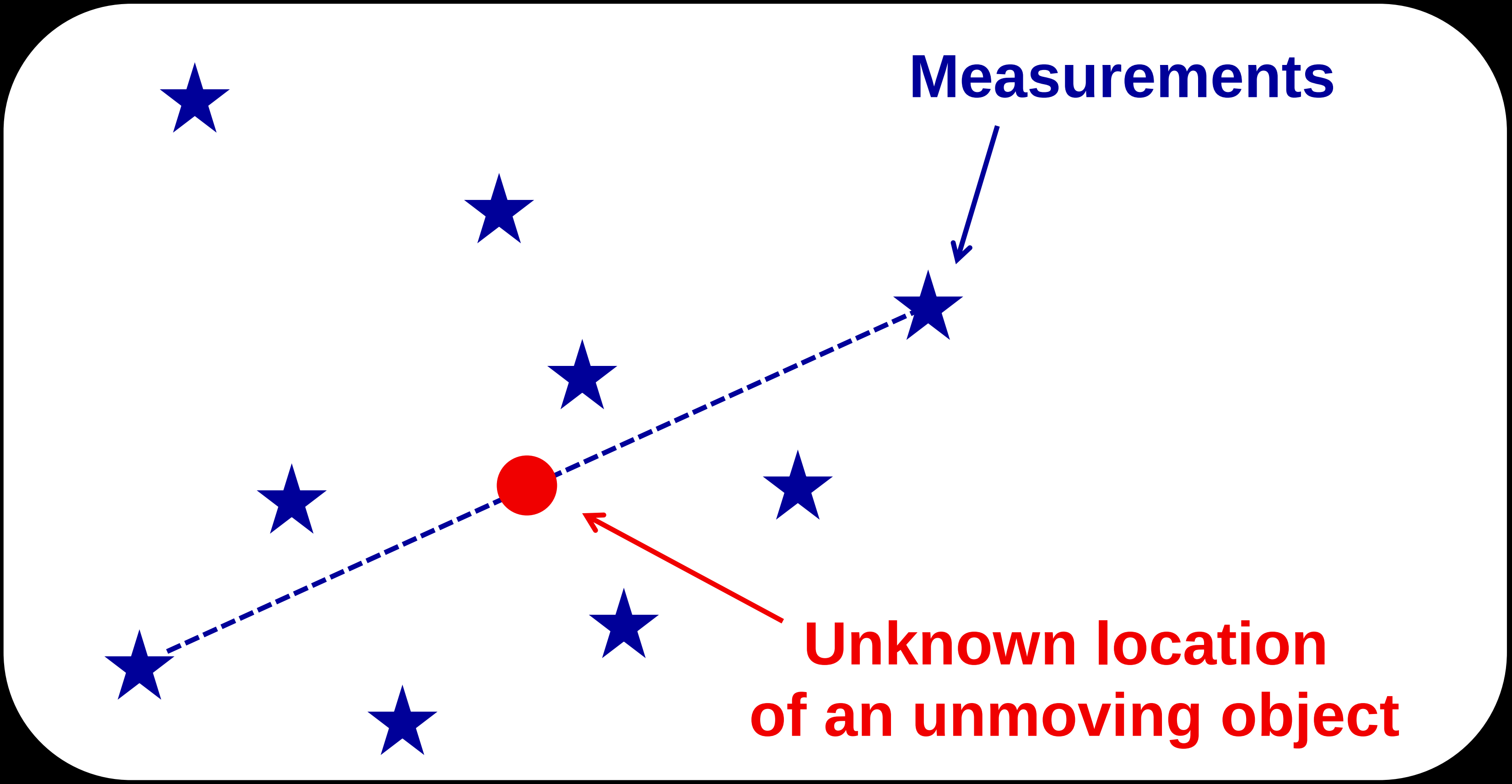
To extract regularities that will allow long-term forecasting we need to smooth data

Which regularities can you extract from measurements of mean arterial pressure?



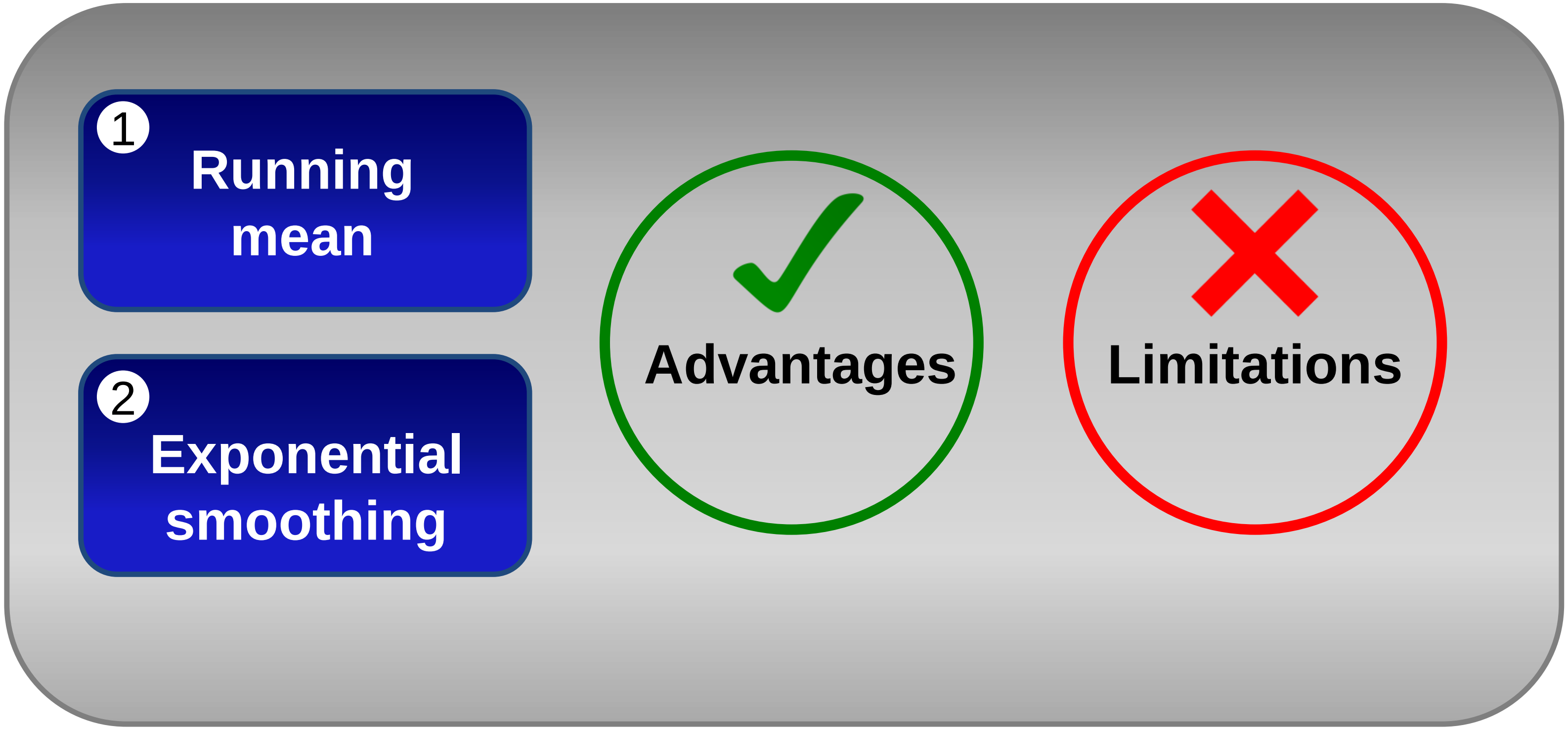
Source of data: Medical Department, Russian University of Friendship of Nations

Estimate the location of an unmoving object



Smoothing is weighted averaging of noisy data.
Fluctuation components are self compensated.

The most popular methods of quasi-optimal estimation



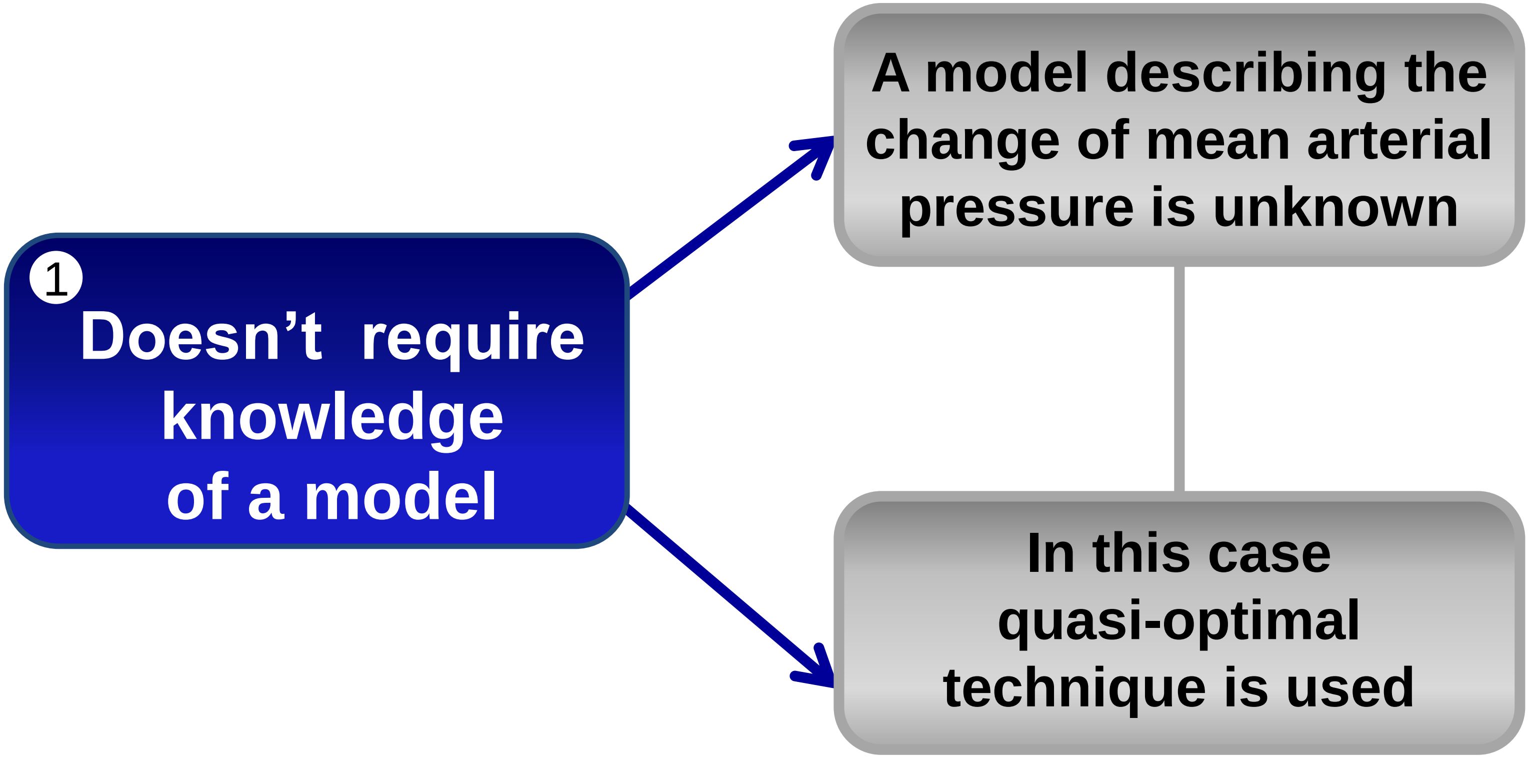
1 Running mean

2 Exponential smoothing

✓ Advantages

✗ Limitations

✓ Advantages of quasi-optimal estimation methods



1

Doesn’t require knowledge of a model

A model describing the change of mean arterial pressure is unknown

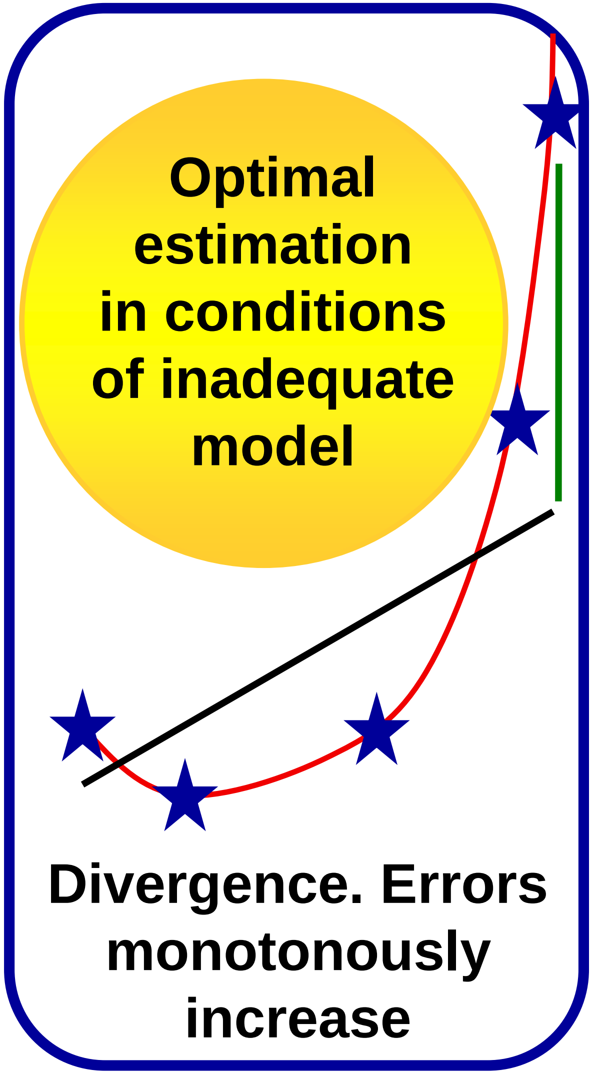
In this case quasi-optimal technique is used

✓ Advantages of quasi-optimal estimation methods

2

Robustness

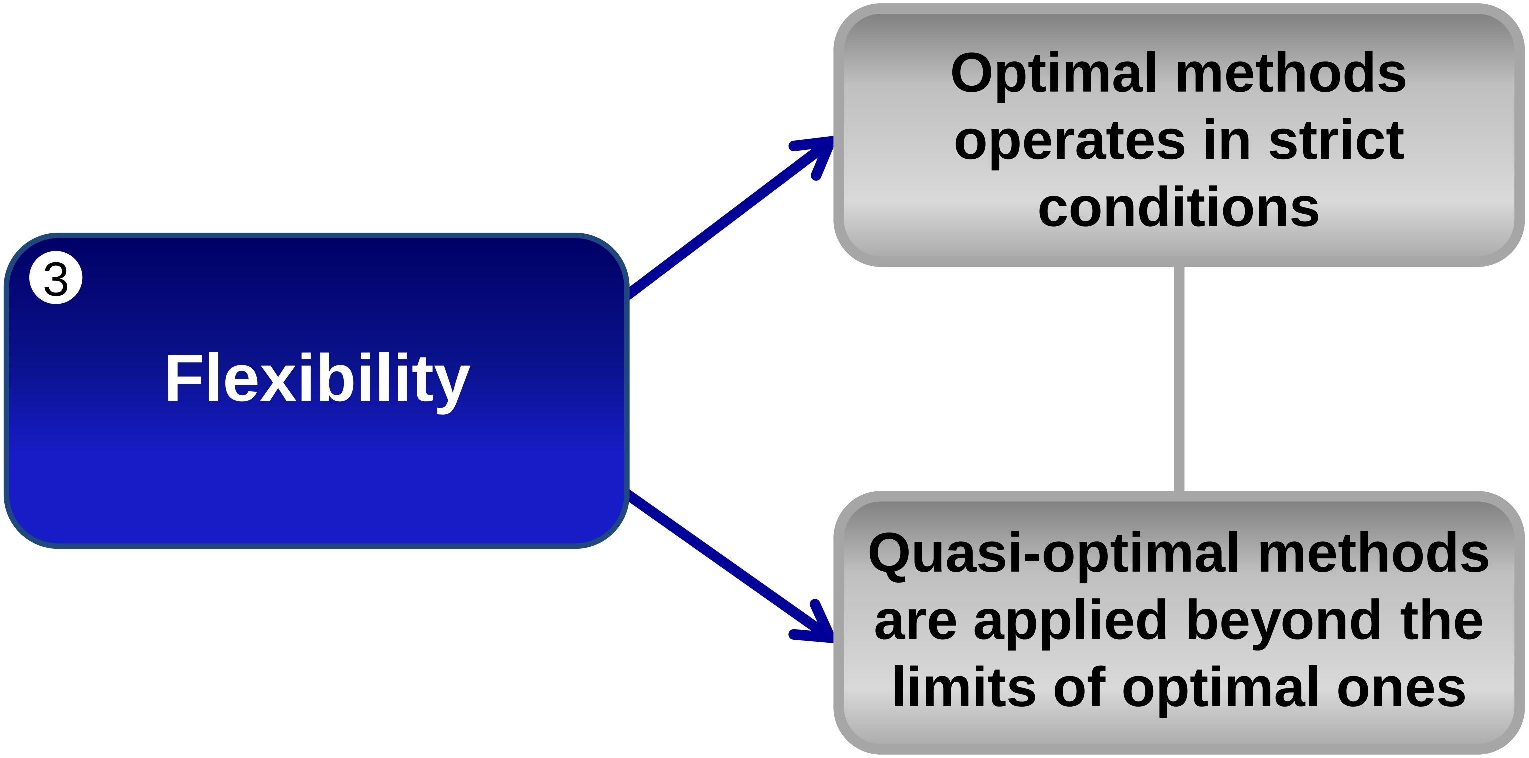
No risk of divergence



Optimal estimation in conditions of inadequate model

Divergence. Errors monotonously increase

✓ Advantages of quasi-optimal estimation methods



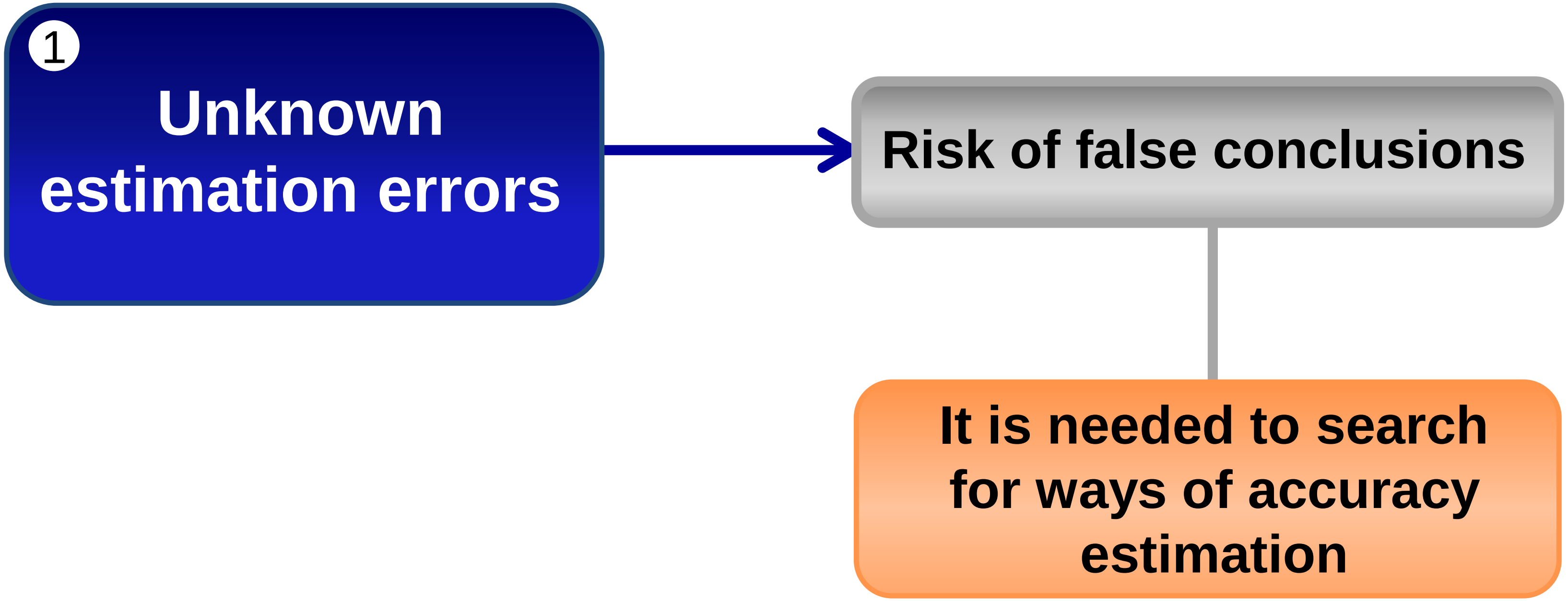
3

Flexibility

Optimal methods operates in strict conditions

Quasi-optimal methods are applied beyond the limits of optimal ones

Disadvantages of quasi-optimal estimation methods

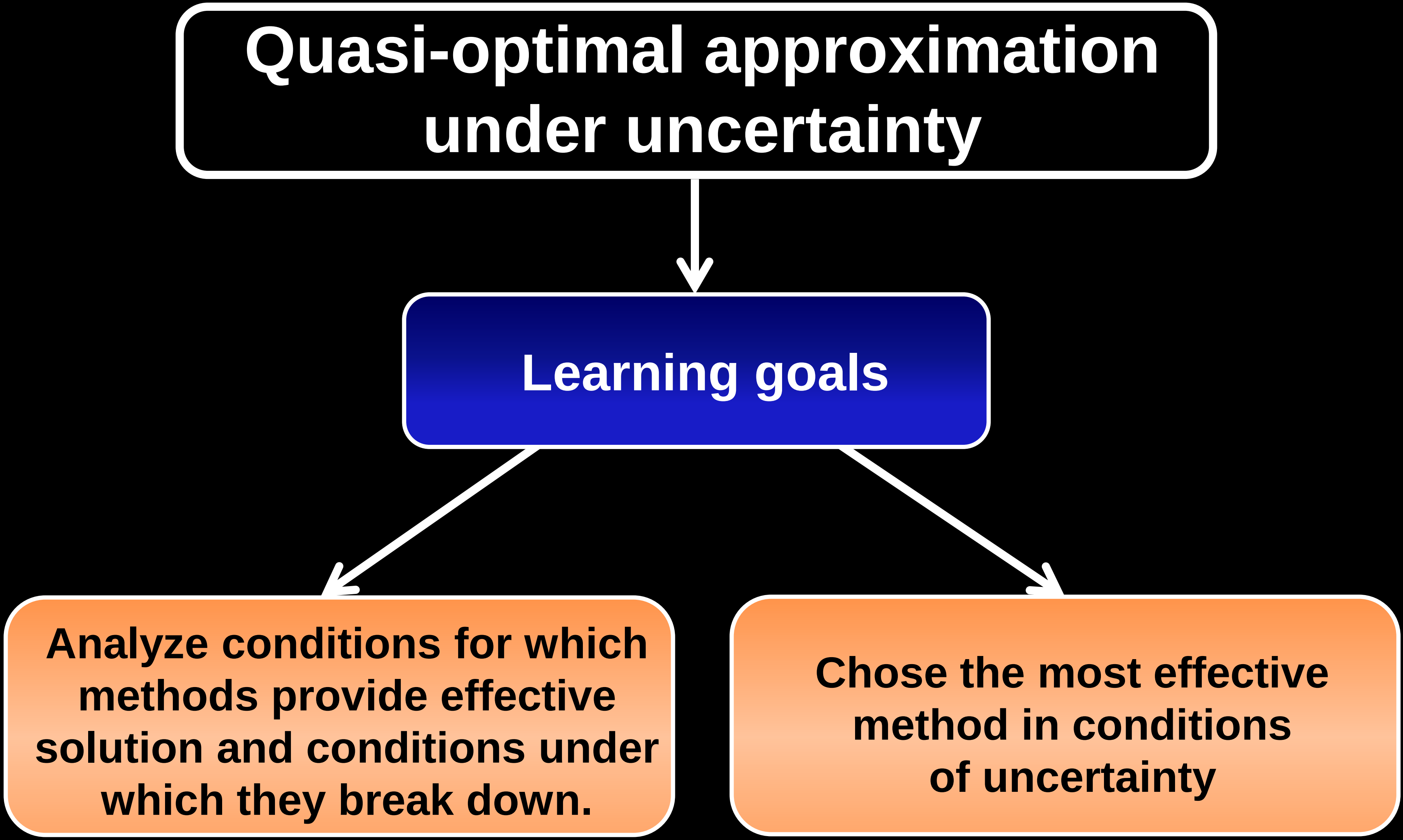


1

Unknown estimation errors

Risk of false conclusions

It is needed to search for ways of accuracy estimation

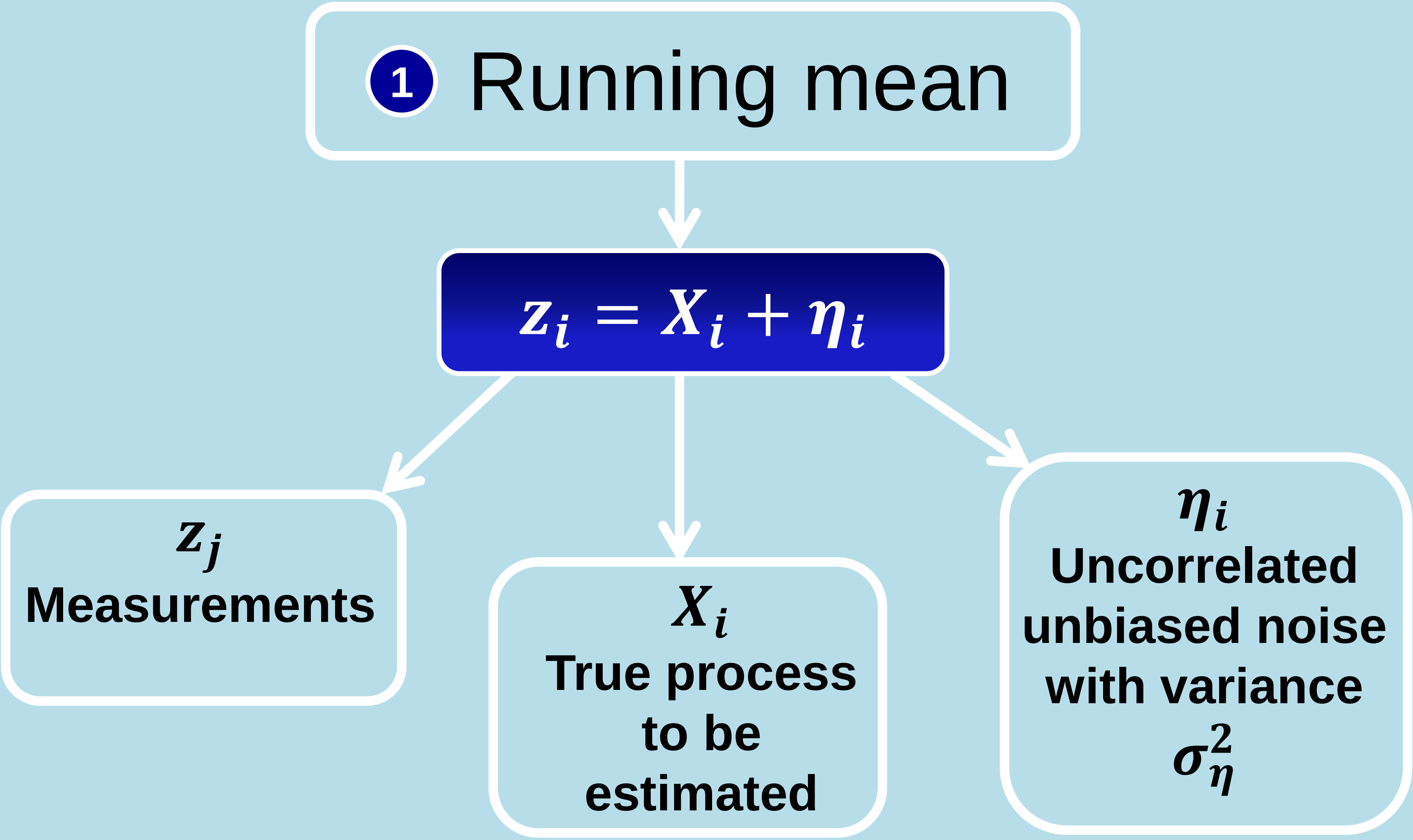


Quasi-optimal approximation under uncertainty

Learning goals

Analyze conditions for which methods provide effective solution and conditions under which they break down.

Chose the most effective method in conditions of uncertainty



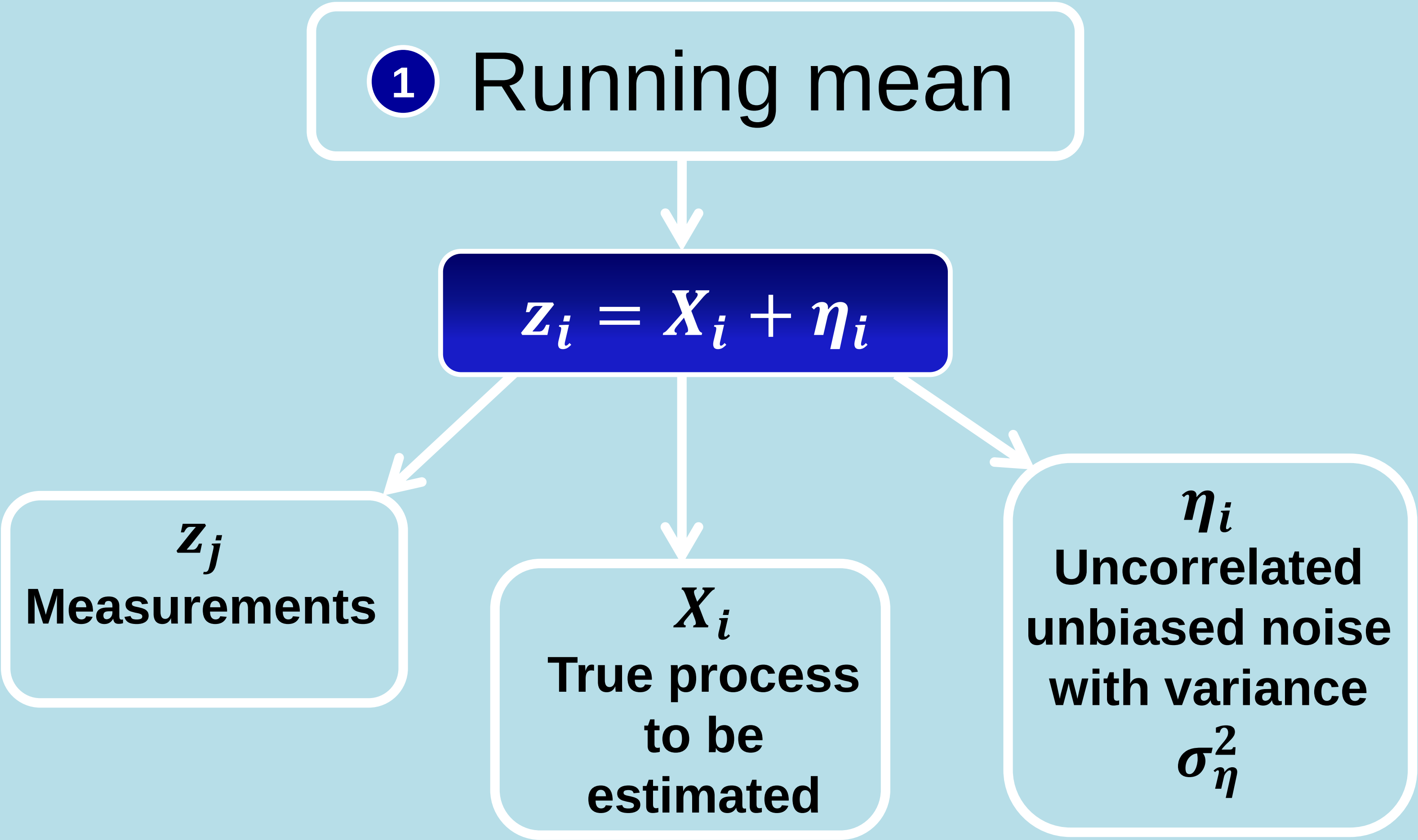
1 Running mean

$$z_i = X_i + \eta_i$$

z_j
Measurements

X_i
True process to be estimated

η_i
Uncorrelated unbiased noise with variance σ_η^2



❶ Running mean

$$z_i = X_i + \eta_i$$

z_j
Measurements

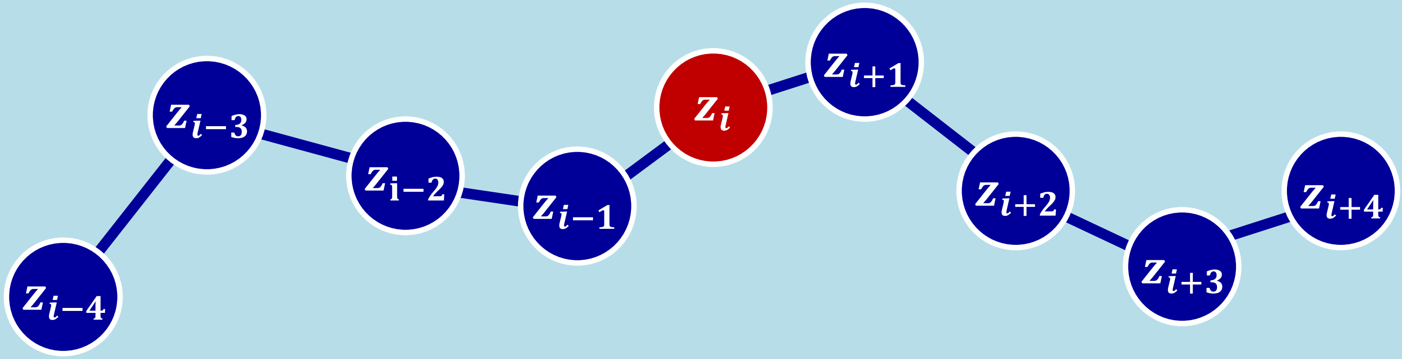
X_i
True process to be estimated

η_i
Uncorrelated unbiased noise with variance σ_η^2

Our goal ⇨ **To reconstruct the dynamics of process X_i using available measurements z_j when the dynamical model is unknown**

① Running mean

Window size $M = 9$

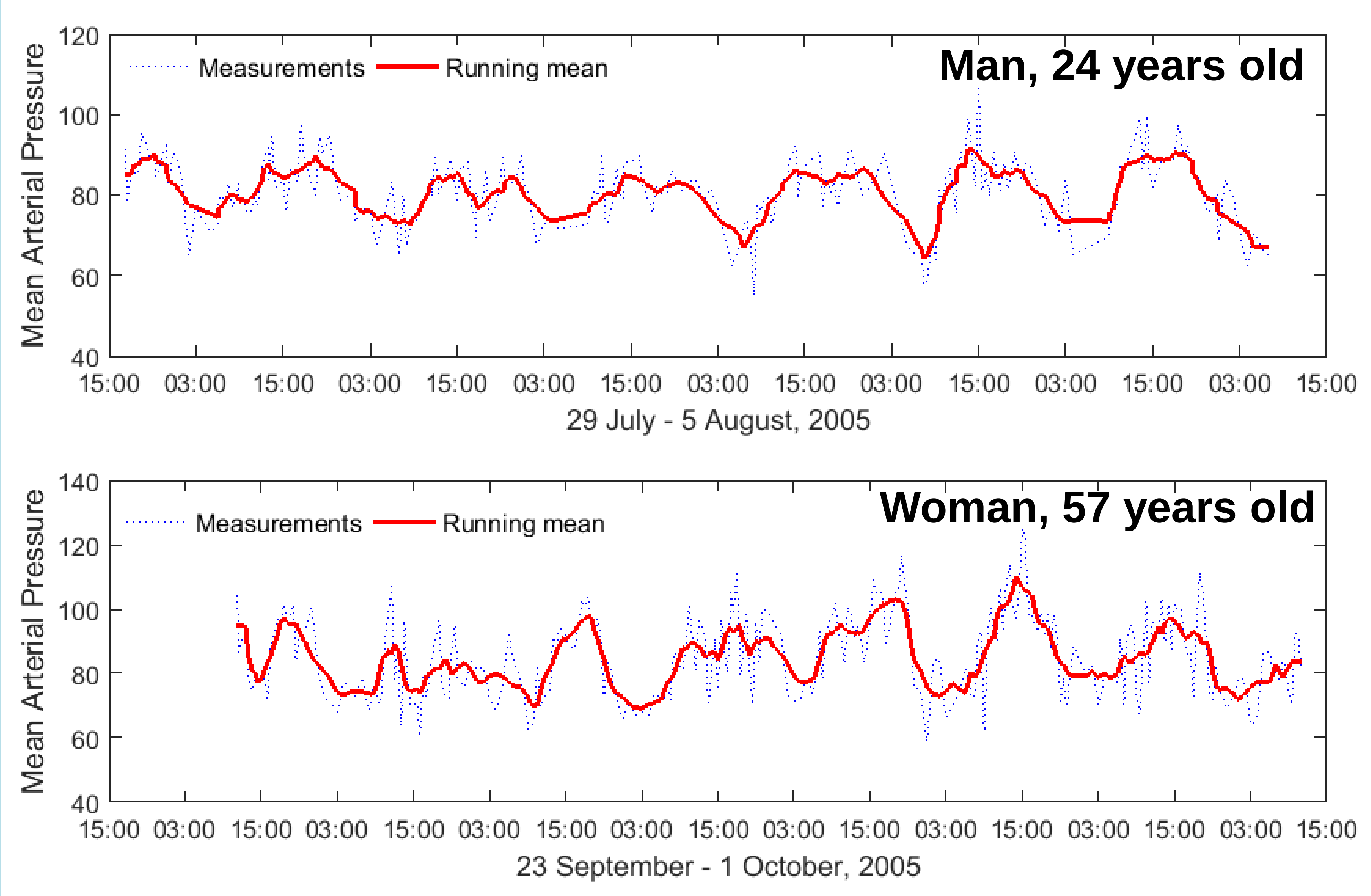


Last 9 measurements z_i

Estimation $\hat{X}_i$ → $\hat{X}_i = \frac{1}{9}\sum_{k=i-4}^{i+4} z_i$ → $\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_i$

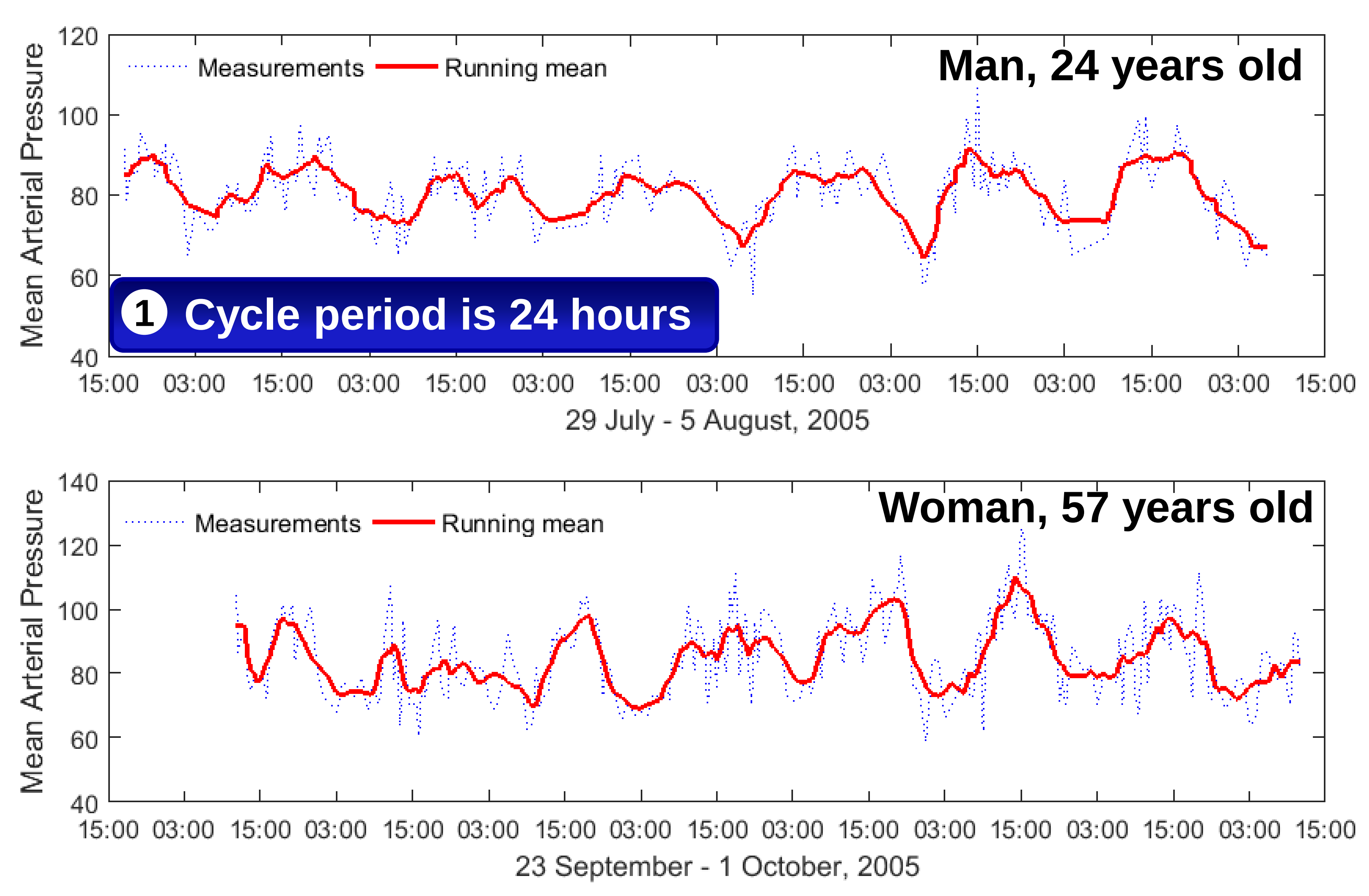
M - window size

Which regularities can you extract from smoothed measurements of mean arterial pressure?



Running mean with window $M = 7$

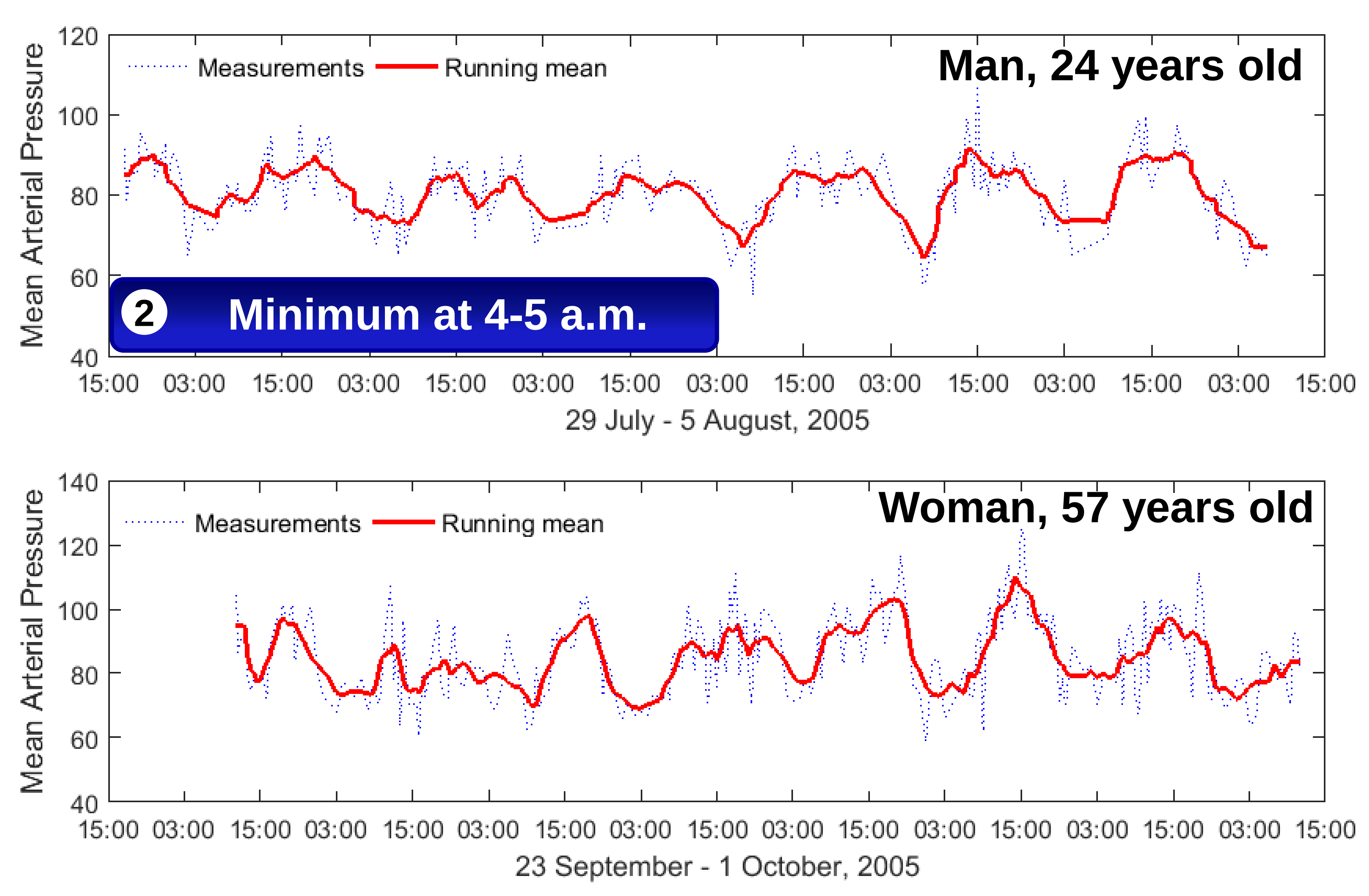
Which regularities can you extract from smoothed measurements of mean arterial pressure?



1 Cycle period is 24 hours

Running mean with window $M = 7$

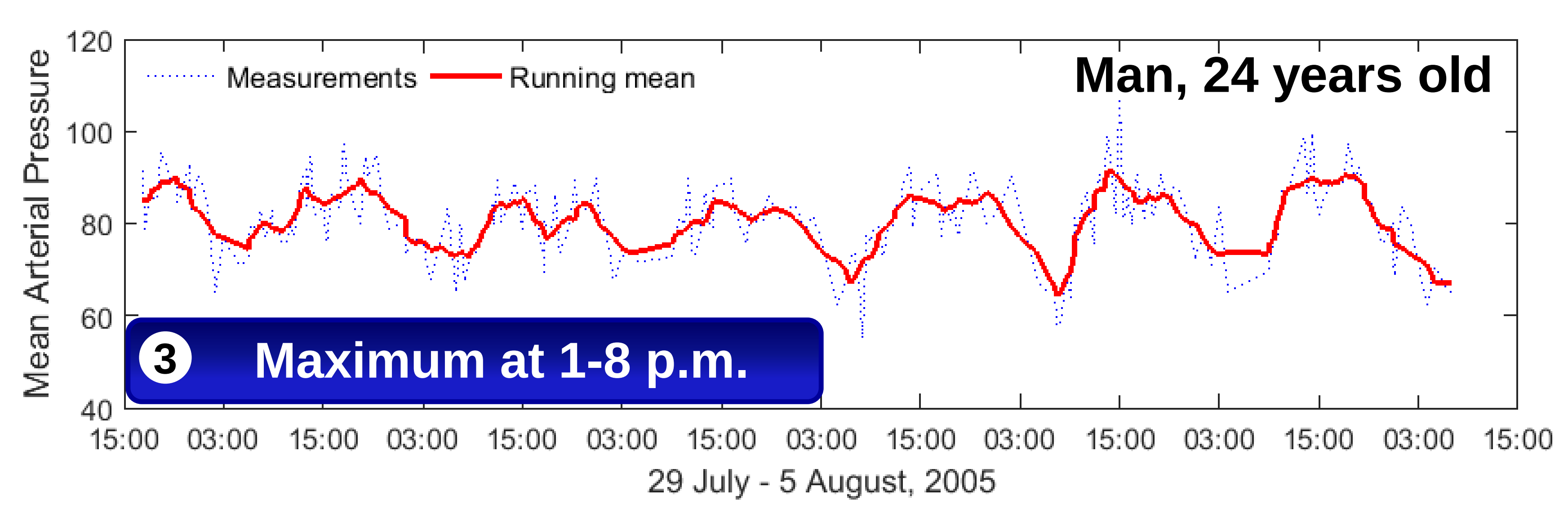
Which regularities can you extract from smoothed measurements of mean arterial pressure?



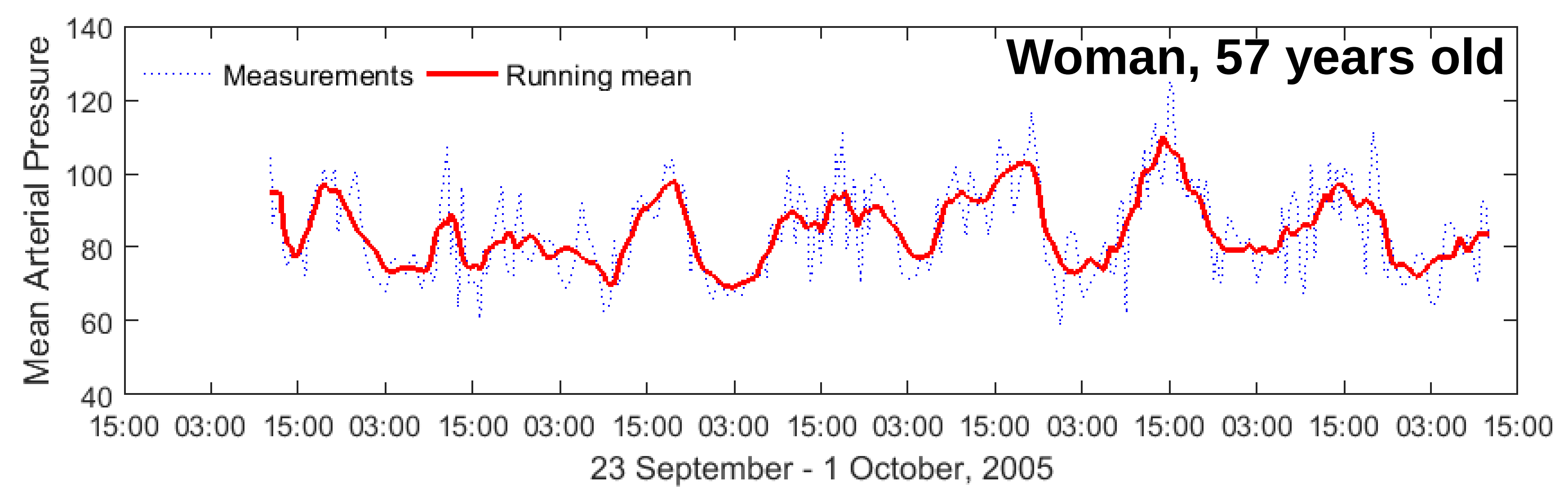
2 Minimum at 4-5 a.m.

Running mean with window $M = 7$

Which regularities can you extract from smoothed measurements of mean arterial pressure?

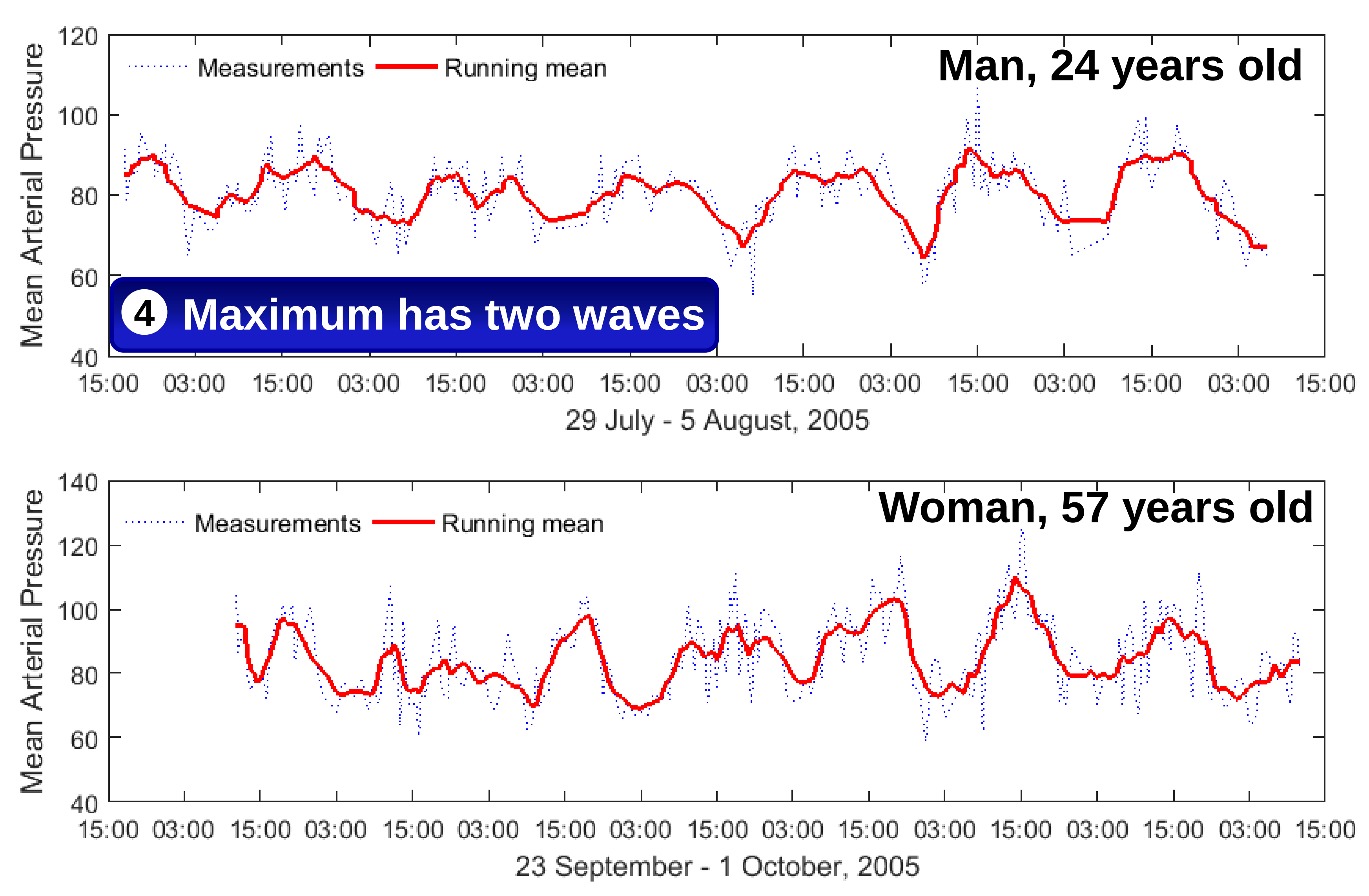


3 **Maximum at 1-8 p.m.**



Running mean with window $M = 7$

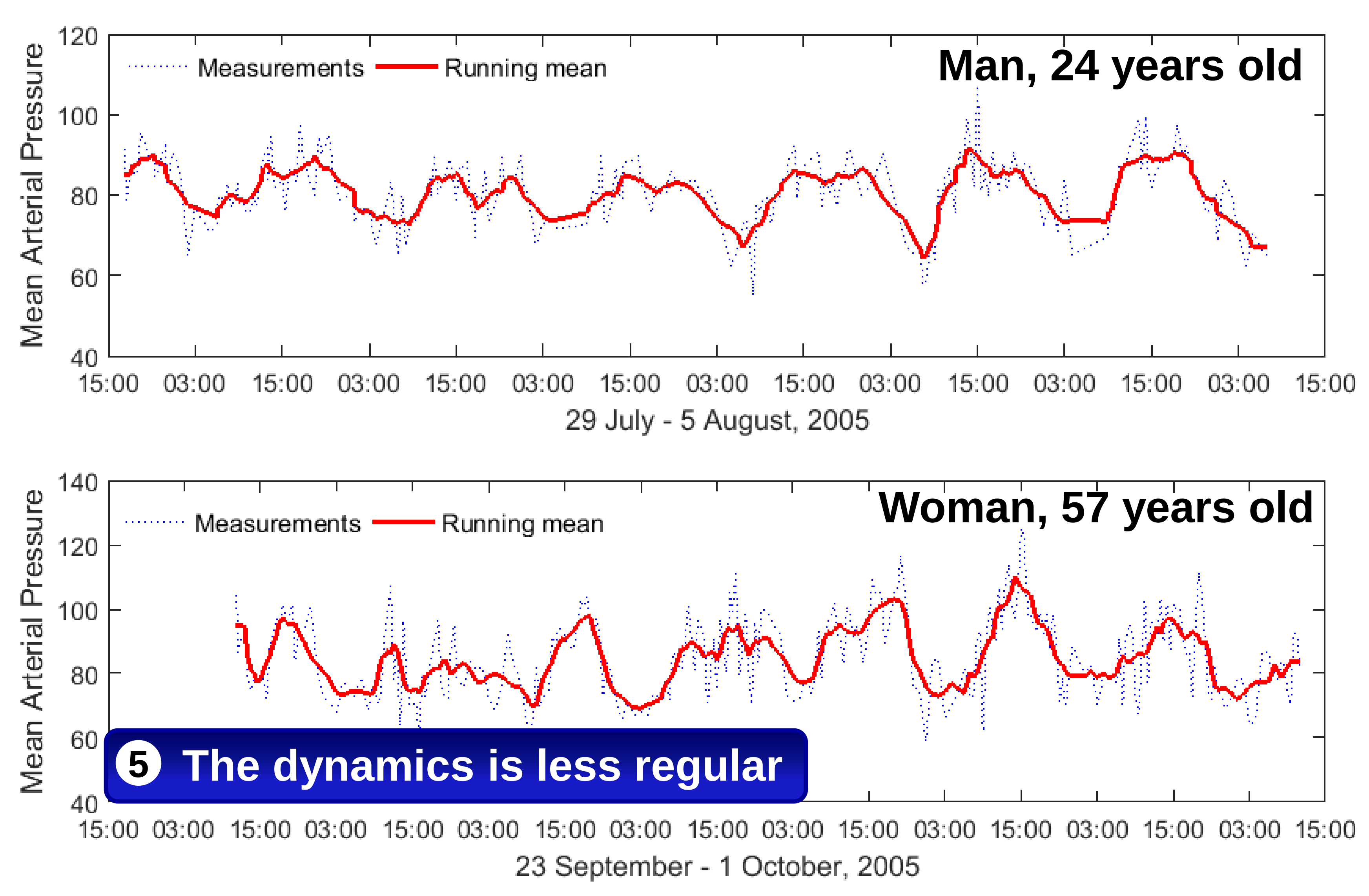
Which regularities can you extract from smoothed measurements of mean arterial pressure?



4 Maximum has two waves

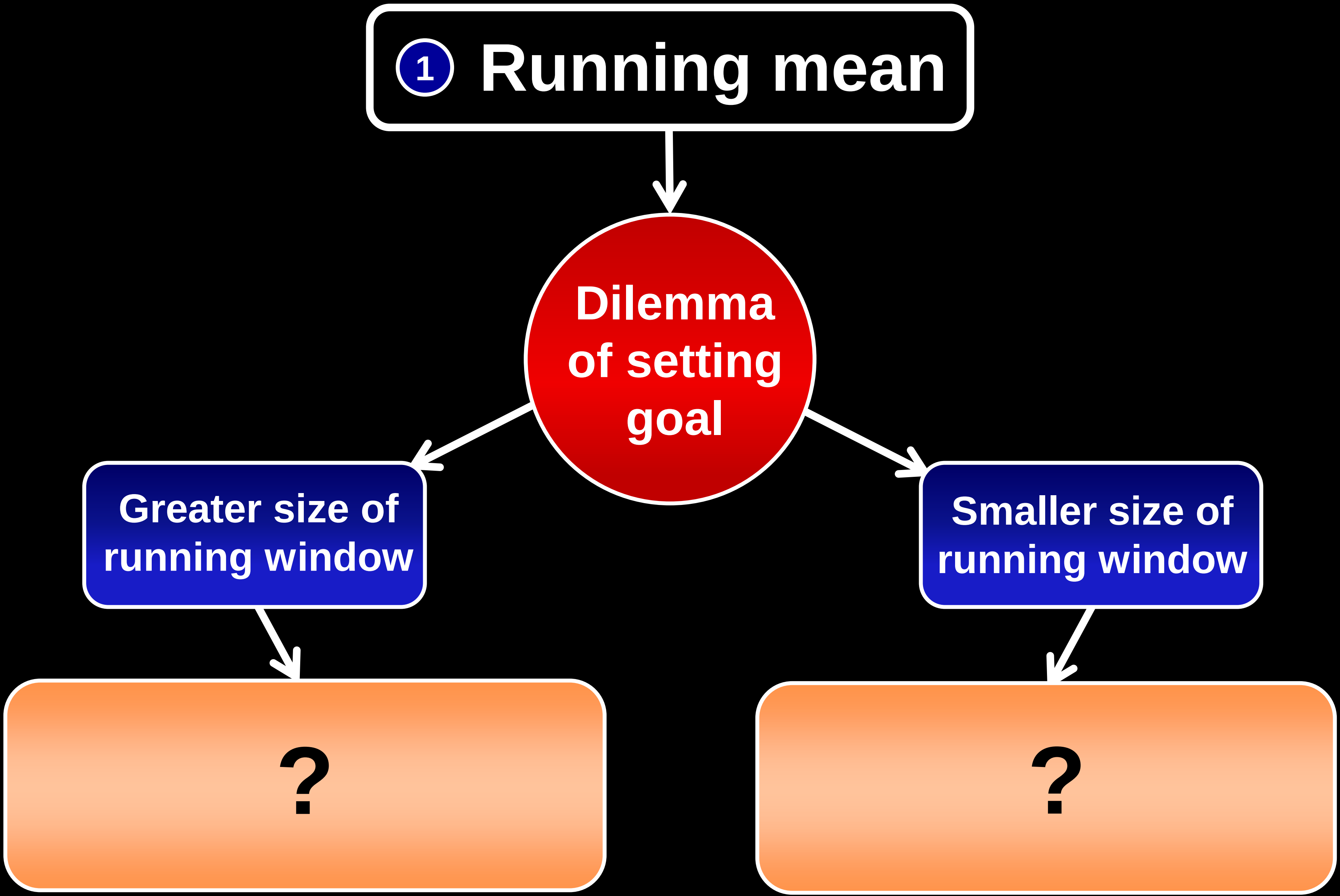
Running mean with window $M = 7$

Which regularities can you extract from smoothed measurements of mean arterial pressure?



5 **The dynamics is less regular**

Running mean with window $M = 7$



1 Running mean

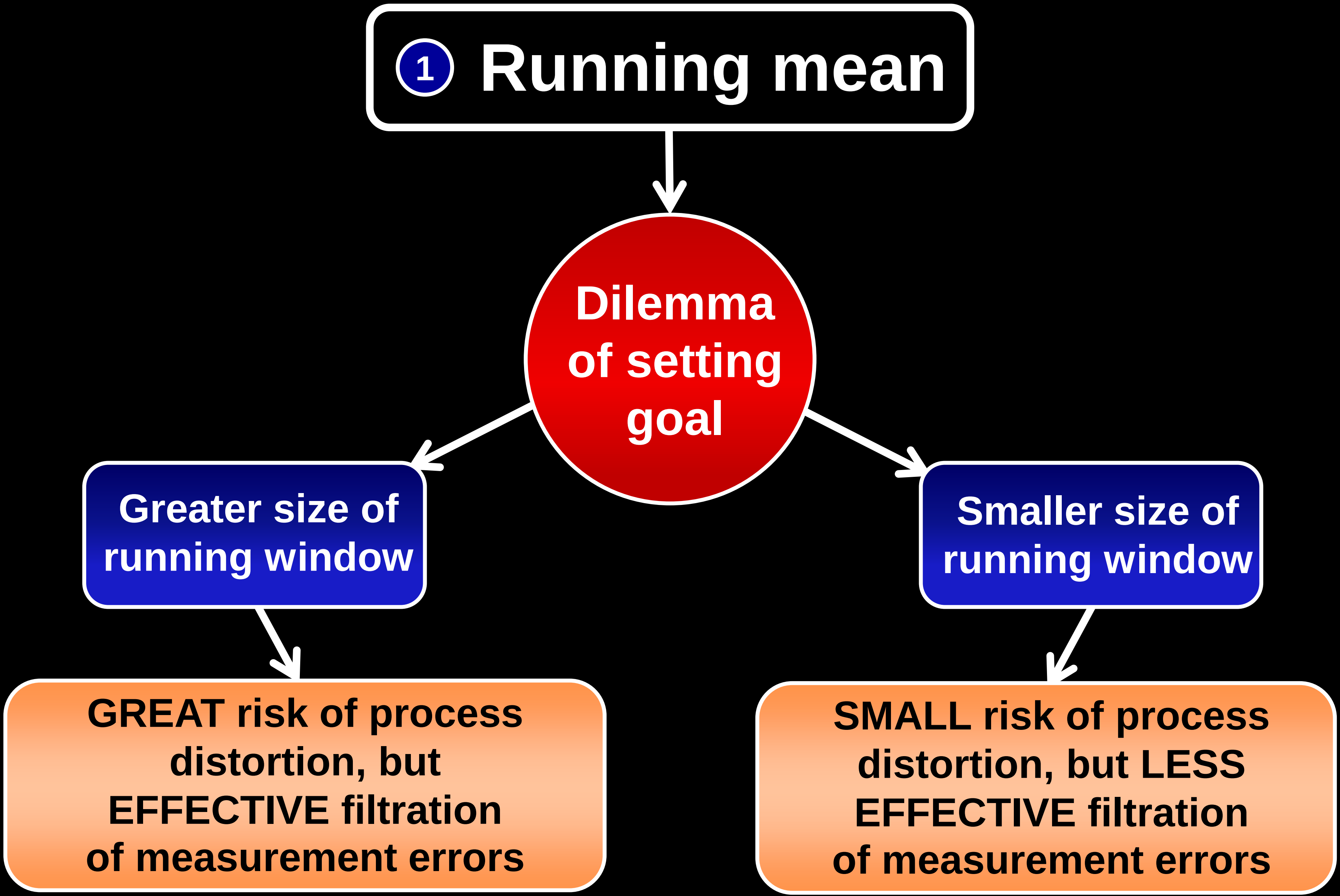
Dilemma of setting goal

Greater size of running window

Smaller size of running window

?

?



1 Running mean

Dilemma of setting goal

Greater size of running window

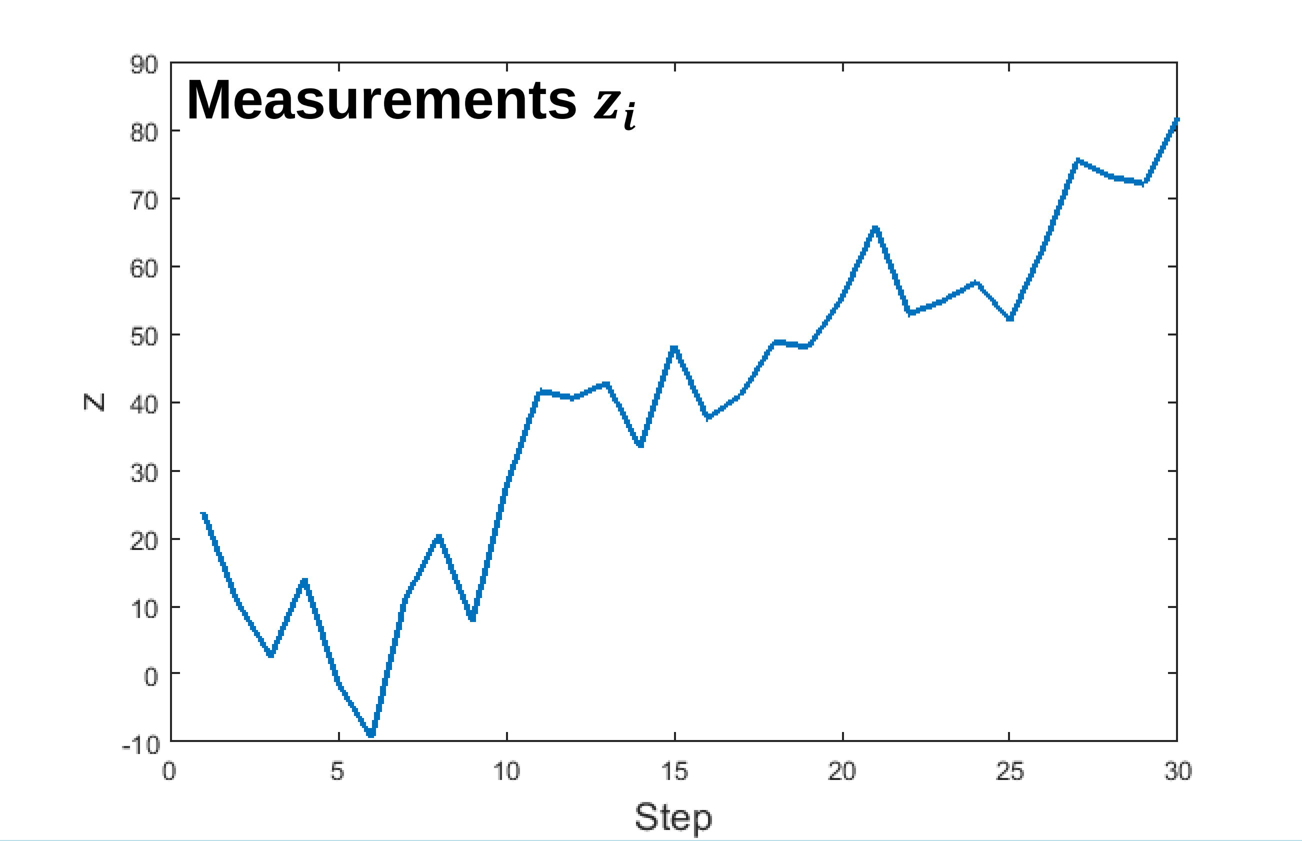
Smaller size of running window

GREAT risk of process distortion, but EFFECTIVE filtration of measurement errors

SMALL risk of process distortion, but LESS EFFECTIVE filtration of measurement errors

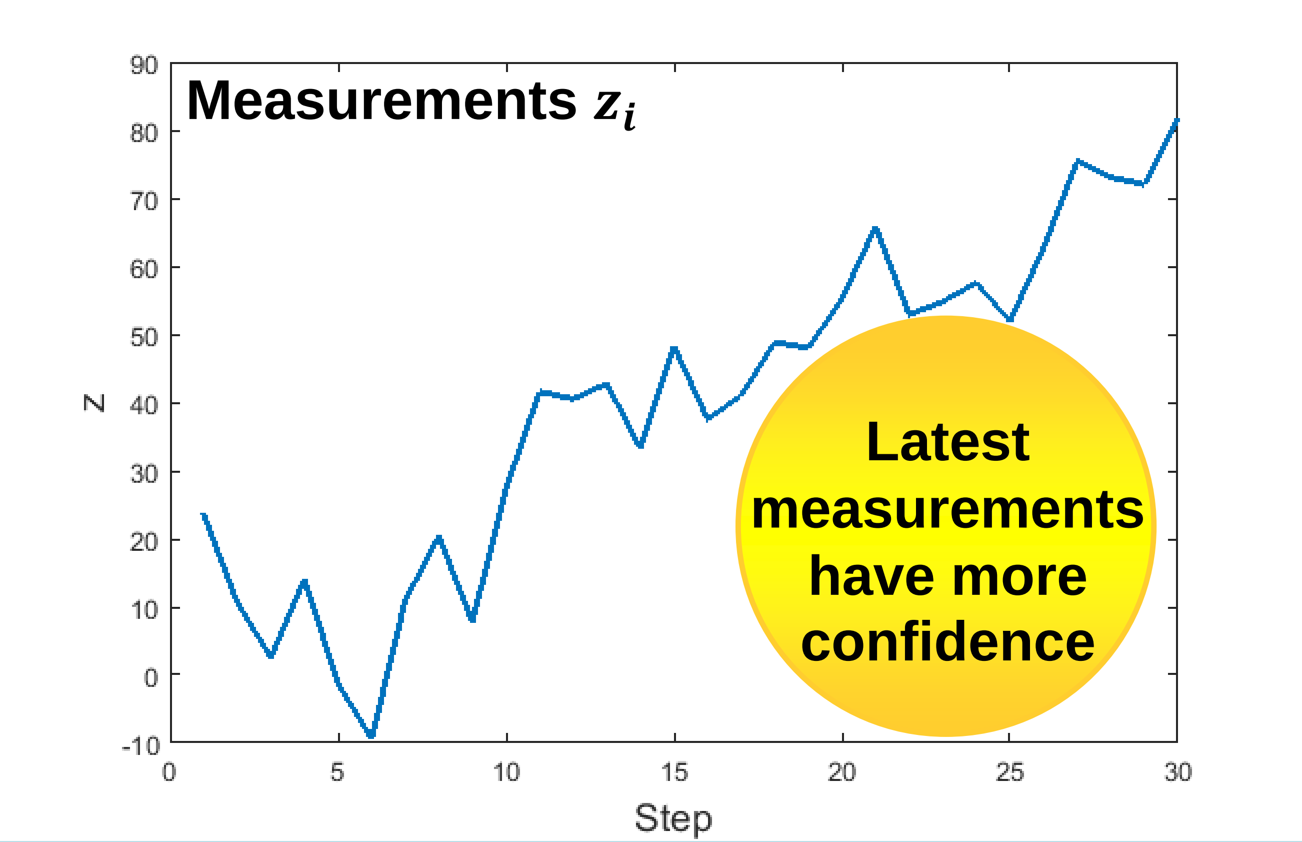
2 Exponential smoothing

Let’s assume that values of process X are characterized by sudden change



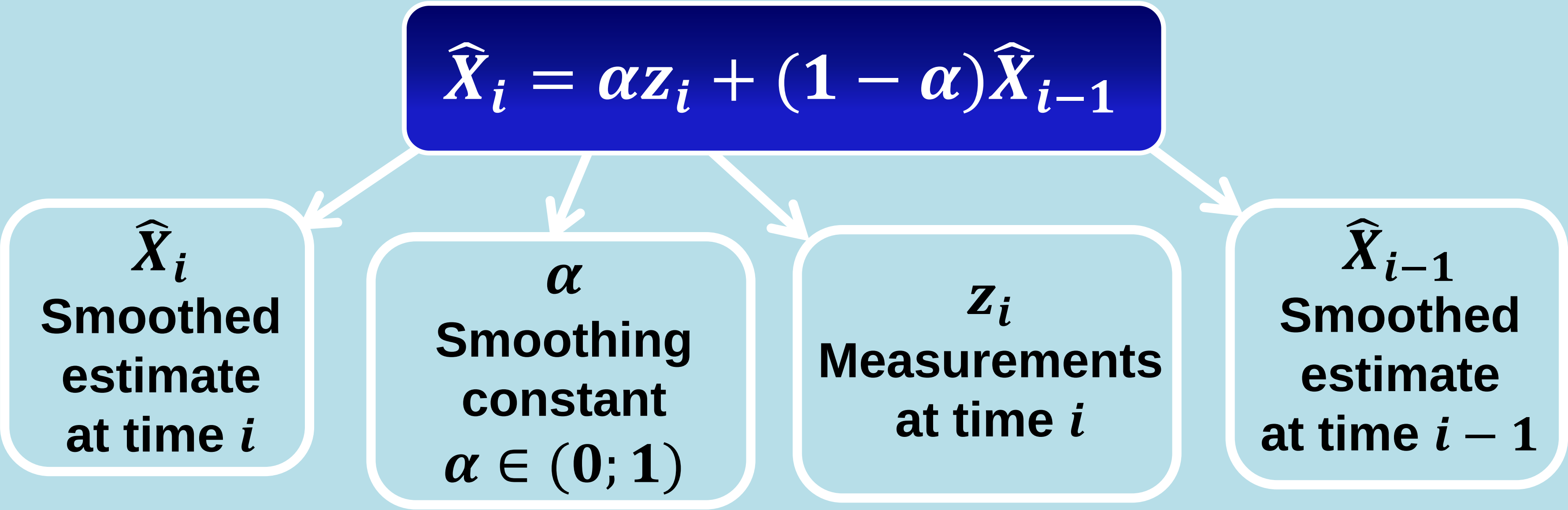
② Exponential smoothing

Let's assume that values of process X are characterized by sudden change



Latest measurements have more confidence

❷ Exponential smoothing



$$\hat{X}_i = \alpha z_i + (1 - \alpha)\hat{X}_{i-1}$$

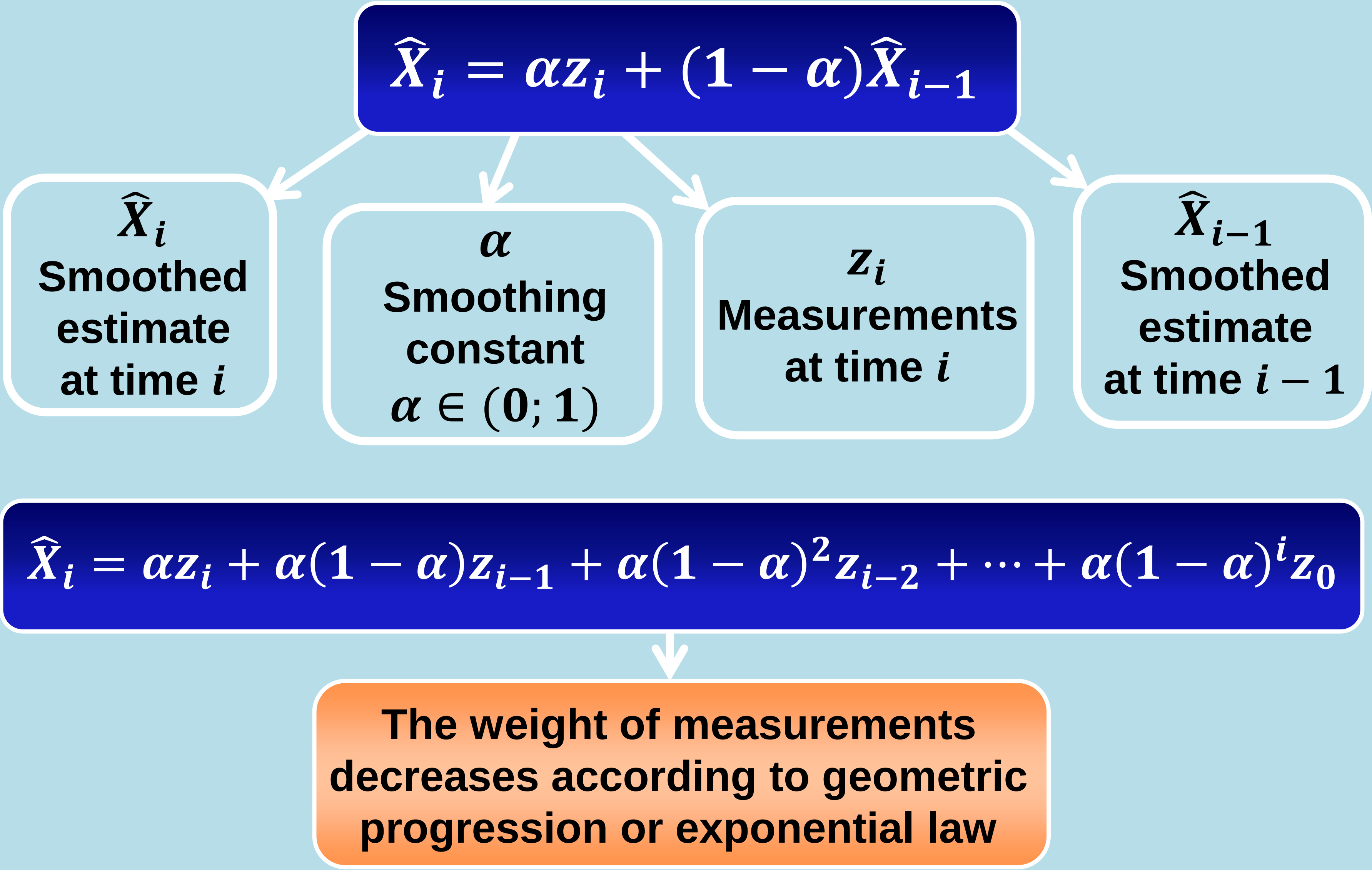
$\hat{X}_i$
Smoothed estimate at time i

α
Smoothing constant
$\alpha \in (0; 1)$

z_i
Measurements at time i

$\hat{X}_{i-1}$
Smoothed estimate at time $i - 1$

② Exponential smoothing



$$\widehat{X}_i = \alpha z_i + (1-\alpha)\widehat{X}_{i-1}$$

$\widehat{X}_i$
Smoothed estimate at time i

α
Smoothing constant
$\alpha \in (0;1)$

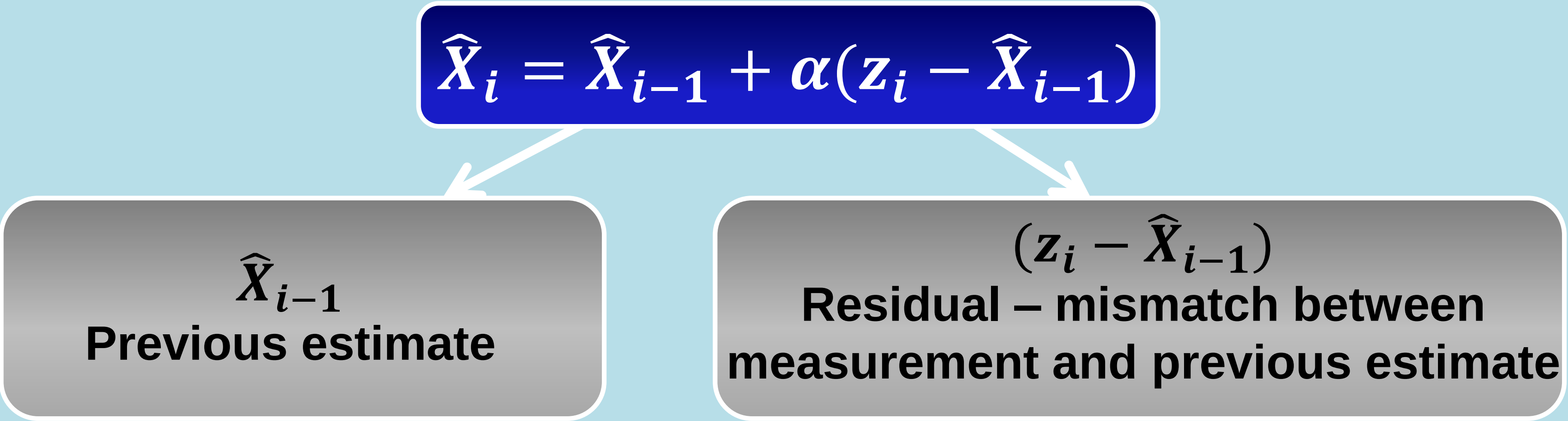
z_i
Measurements at time i

$\widehat{X}_{i-1}$
Smoothed estimate at time $i-1$

$$\widehat{X}_i = \alpha z_i + \alpha(1-\alpha)z_{i-1} + \alpha(1-\alpha)^2 z_{i-2} + \cdots + \alpha(1-\alpha)^i z_0$$

The weight of measurements decreases according to geometric progression or exponential law

❷ Exponential smoothing: Dilemma of setting goal



$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

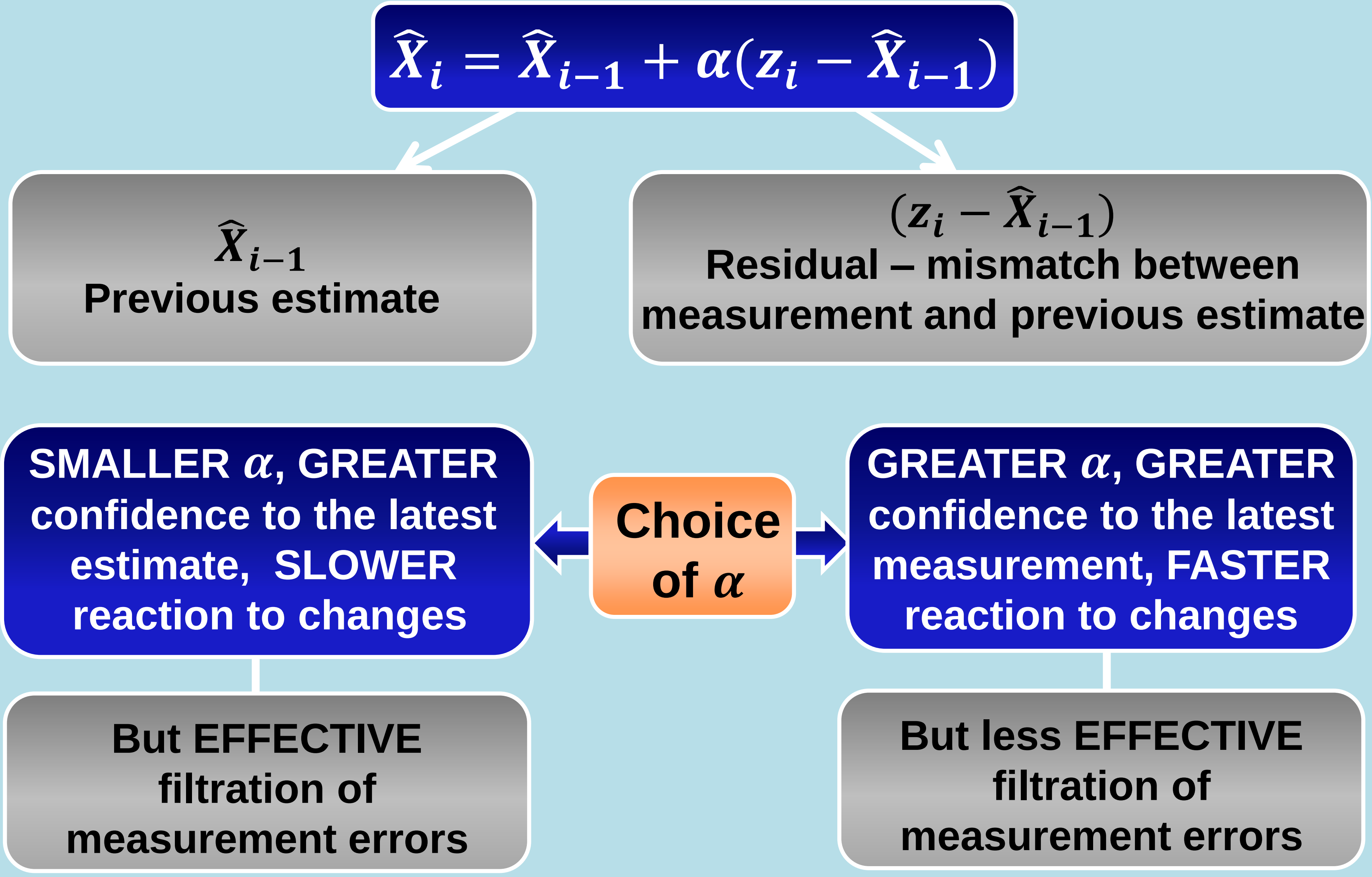
$\hat{X}_{i-1}$

Previous estimate

$(z_i - \hat{X}_{i-1})$

Residual – mismatch between measurement and previous estimate

❷ Exponential smoothing: Dilemma of setting goal



$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

$\hat{X}_{i-1}$

Previous estimate

$(z_i - \hat{X}_{i-1})$

Residual – mismatch between measurement and previous estimate

SMALLER α, GREATER confidence to the latest estimate, SLOWER reaction to changes

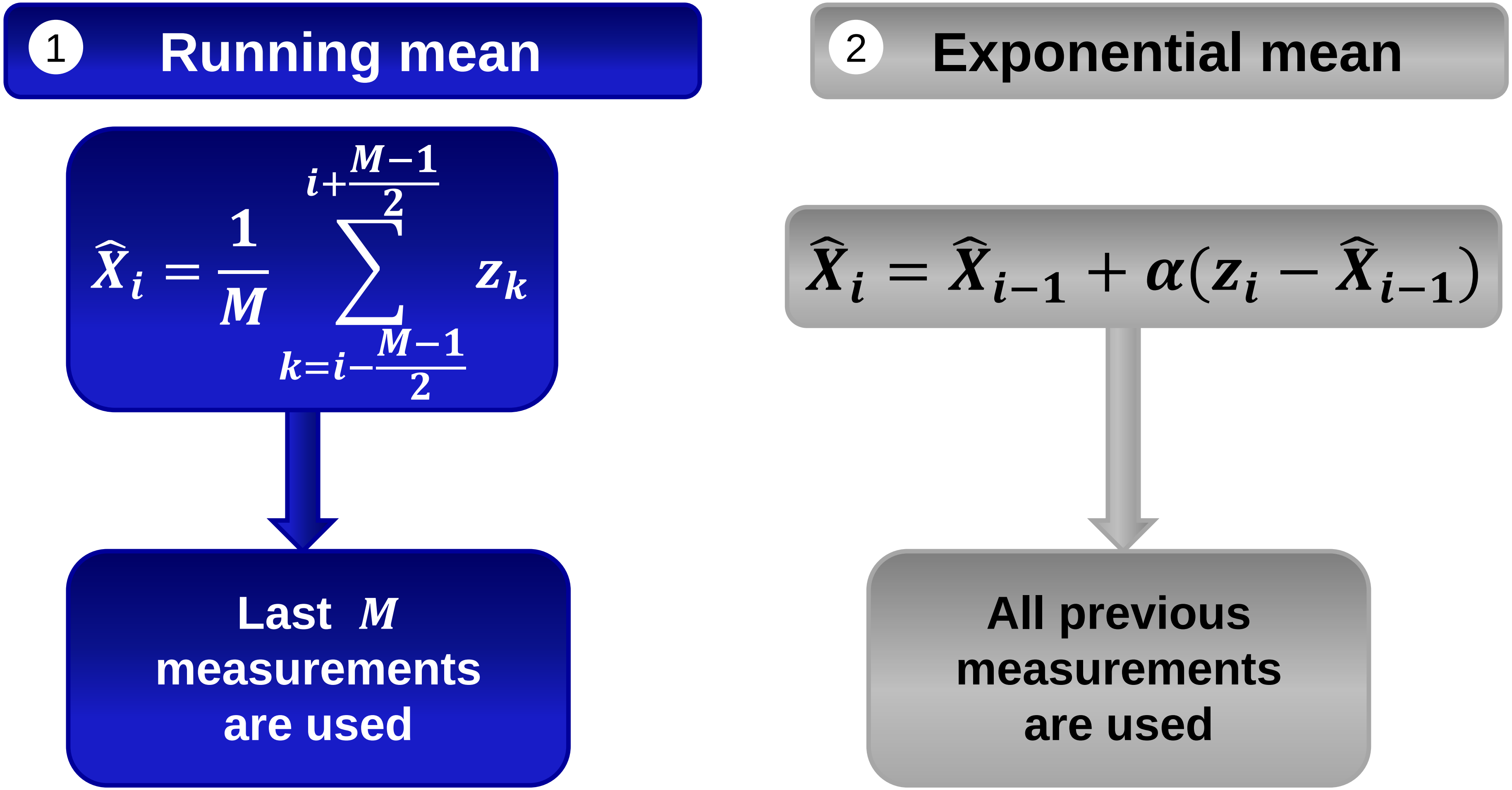
Choice of α

GREATER α, GREATER confidence to the latest measurement, FASTER reaction to changes

But EFFECTIVE filtration of measurement errors

But less EFFECTIVE filtration of measurement errors

Comparison of smoothing methods



1 **Running mean**

$$\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

Last M measurements are used

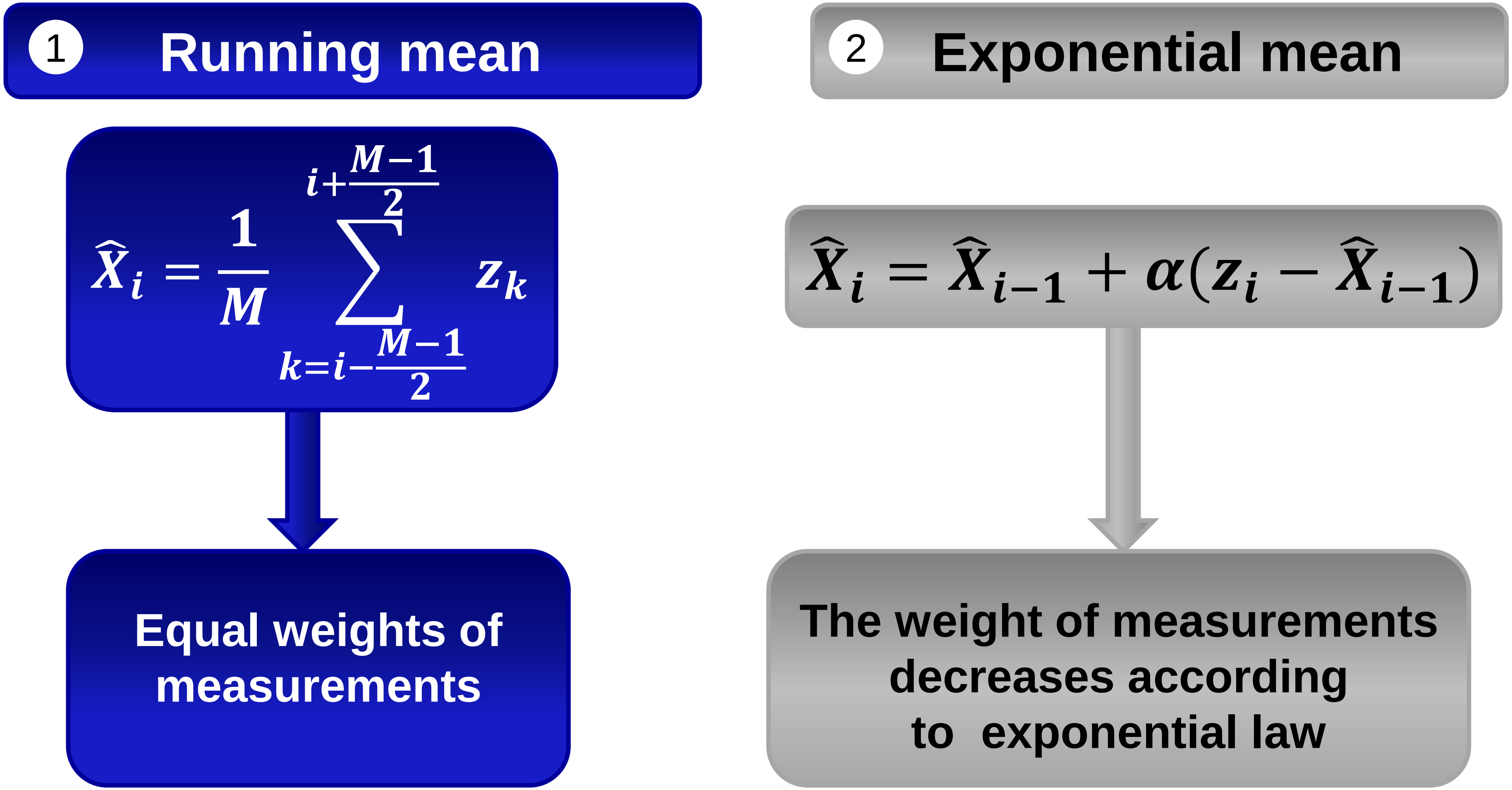
2 **Exponential mean**

$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

All previous measurements are used

1

Comparison of smoothing methods



1 Running mean

$$\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

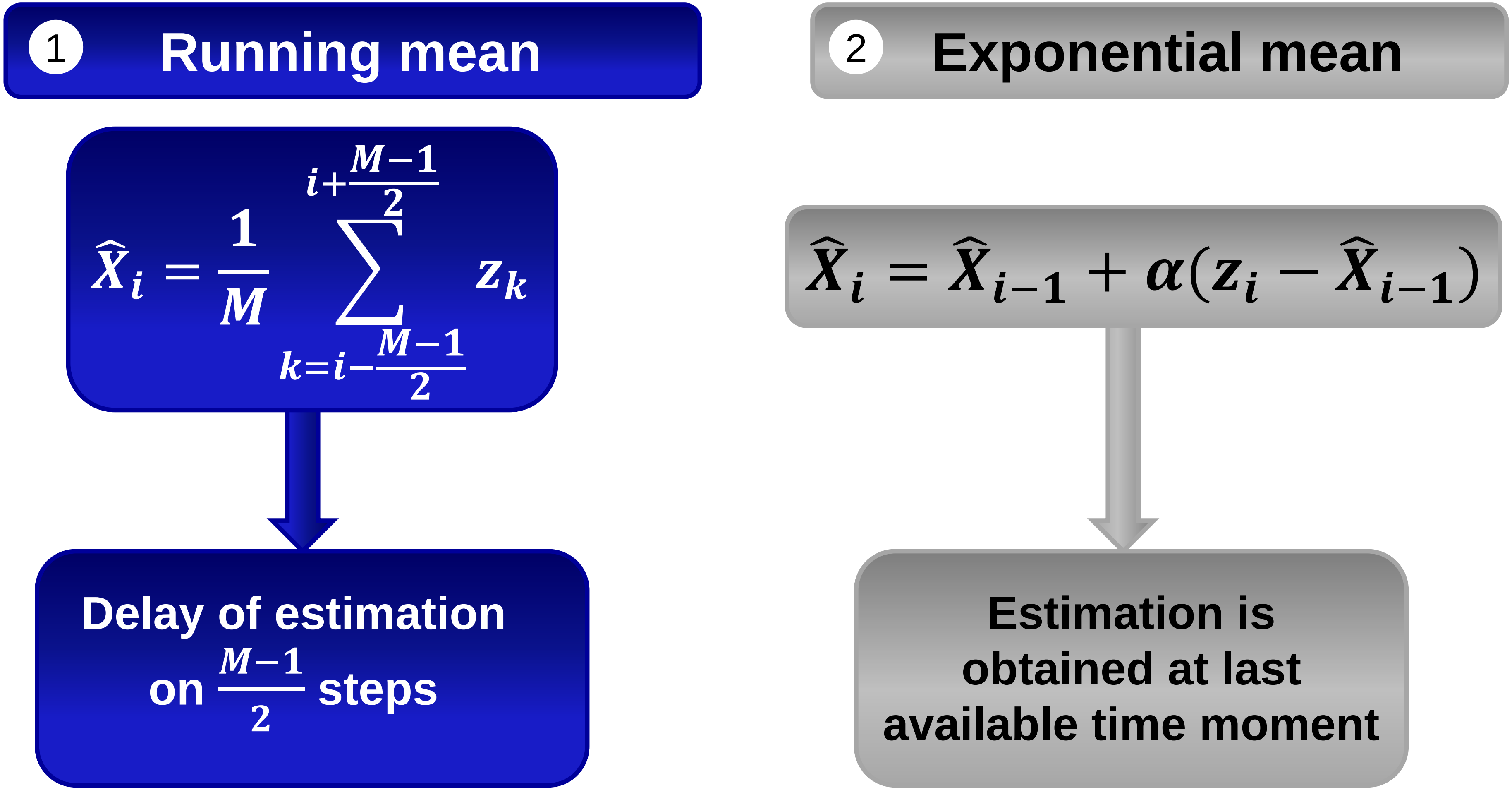
Equal weights of measurements

2 Exponential mean

$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

The weight of measurements decreases according to exponential law

Comparison of smoothing methods



1 Running mean

$$\hat{X}_i = \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

Delay of estimation on $\frac{M-1}{2}$ steps

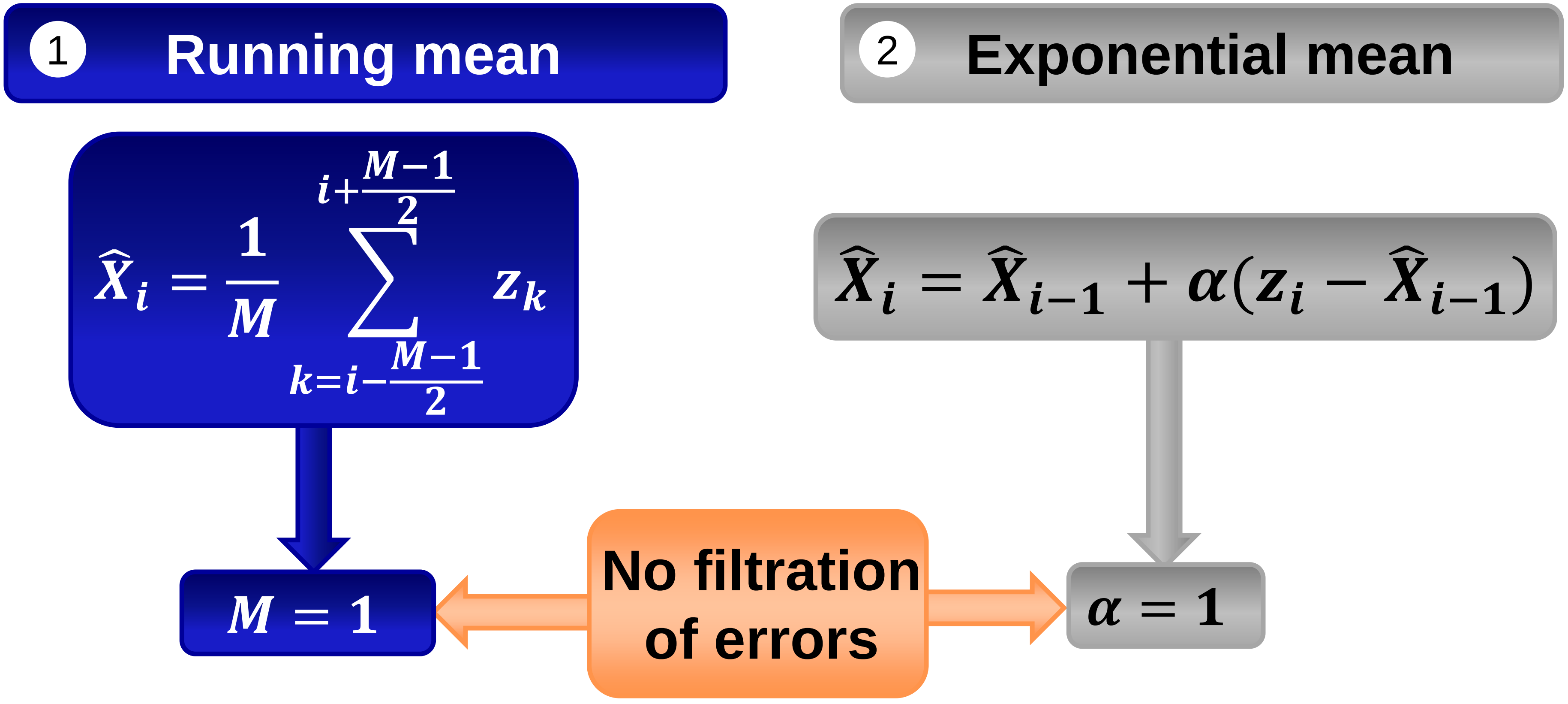
2 Exponential mean

$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

Estimation is obtained at last available time moment

3

Comparison of smoothing methods



1 Running mean

$$\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

$$M = 1$$

No filtration of errors

2 Exponential mean

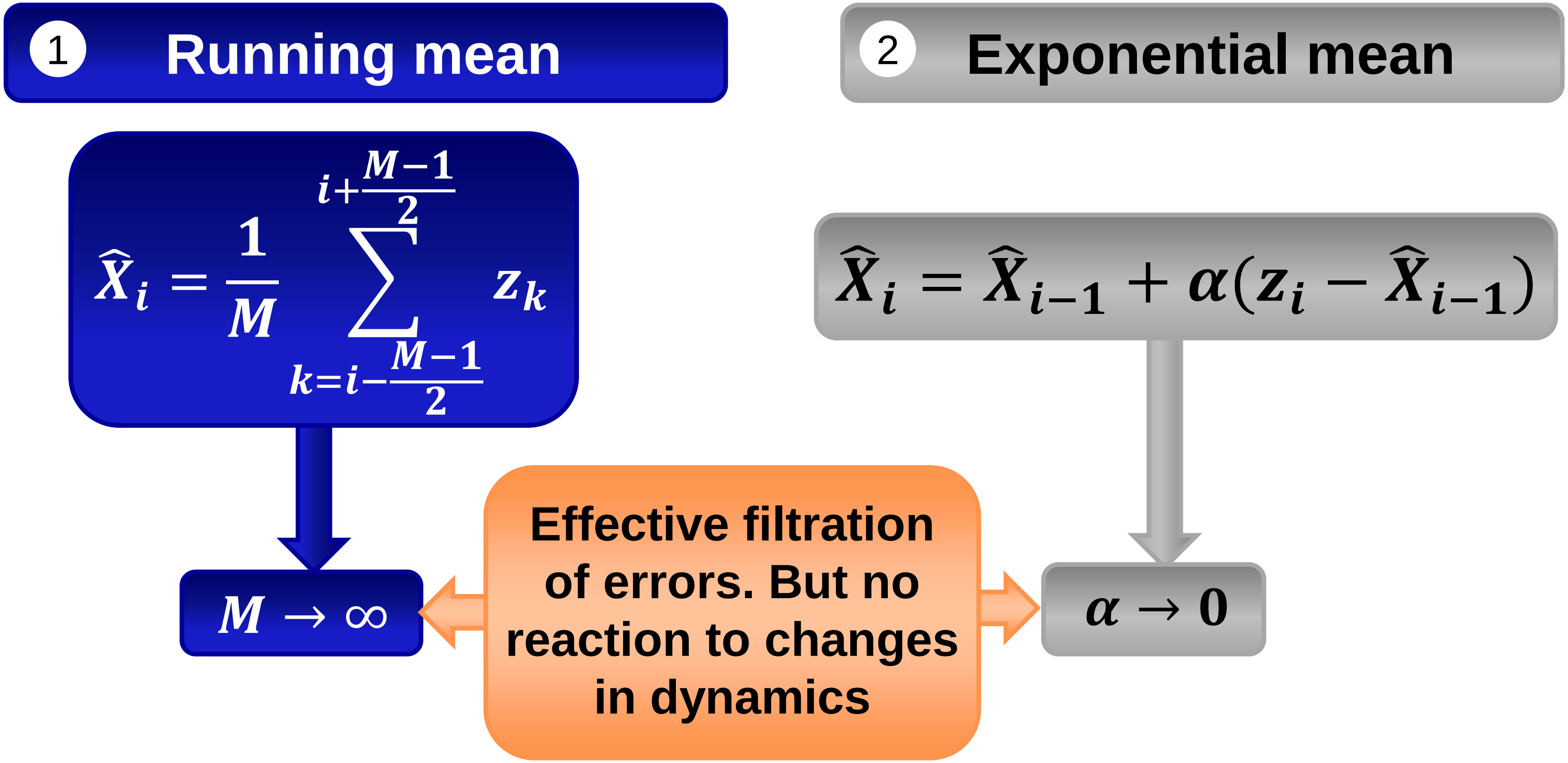
$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

$$\alpha = 1$$

4

Estimates of both smoothing methods are the same

Comparison of smoothing methods



1 Running mean

$$\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

$M \to \infty$

2 Exponential mean

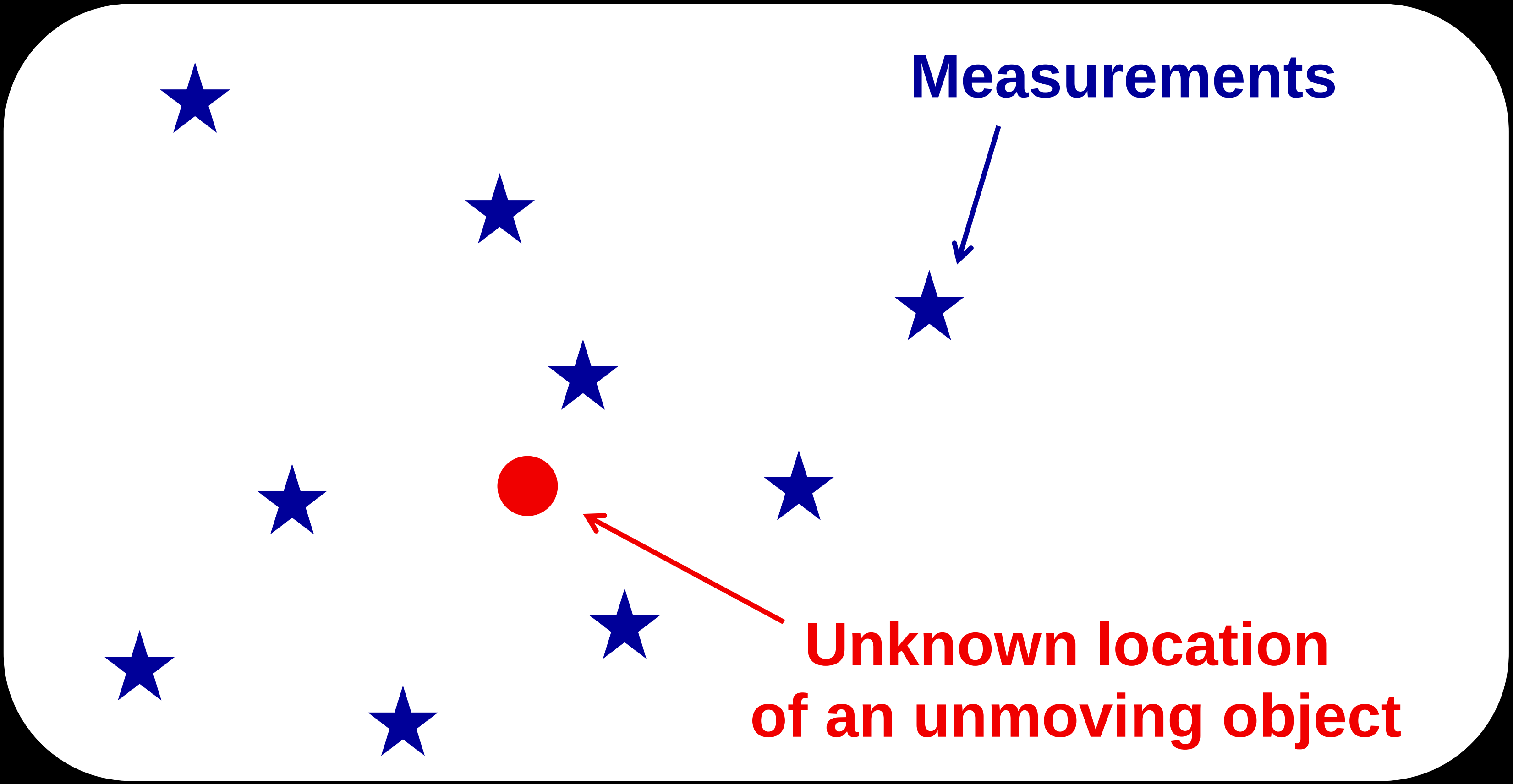
$$\hat{X}_i = \hat{X}_{i-1} + \alpha(z_i - \hat{X}_{i-1})$$

$\alpha \to 0$

Effective filtration of errors. But no reaction to changes in dynamics

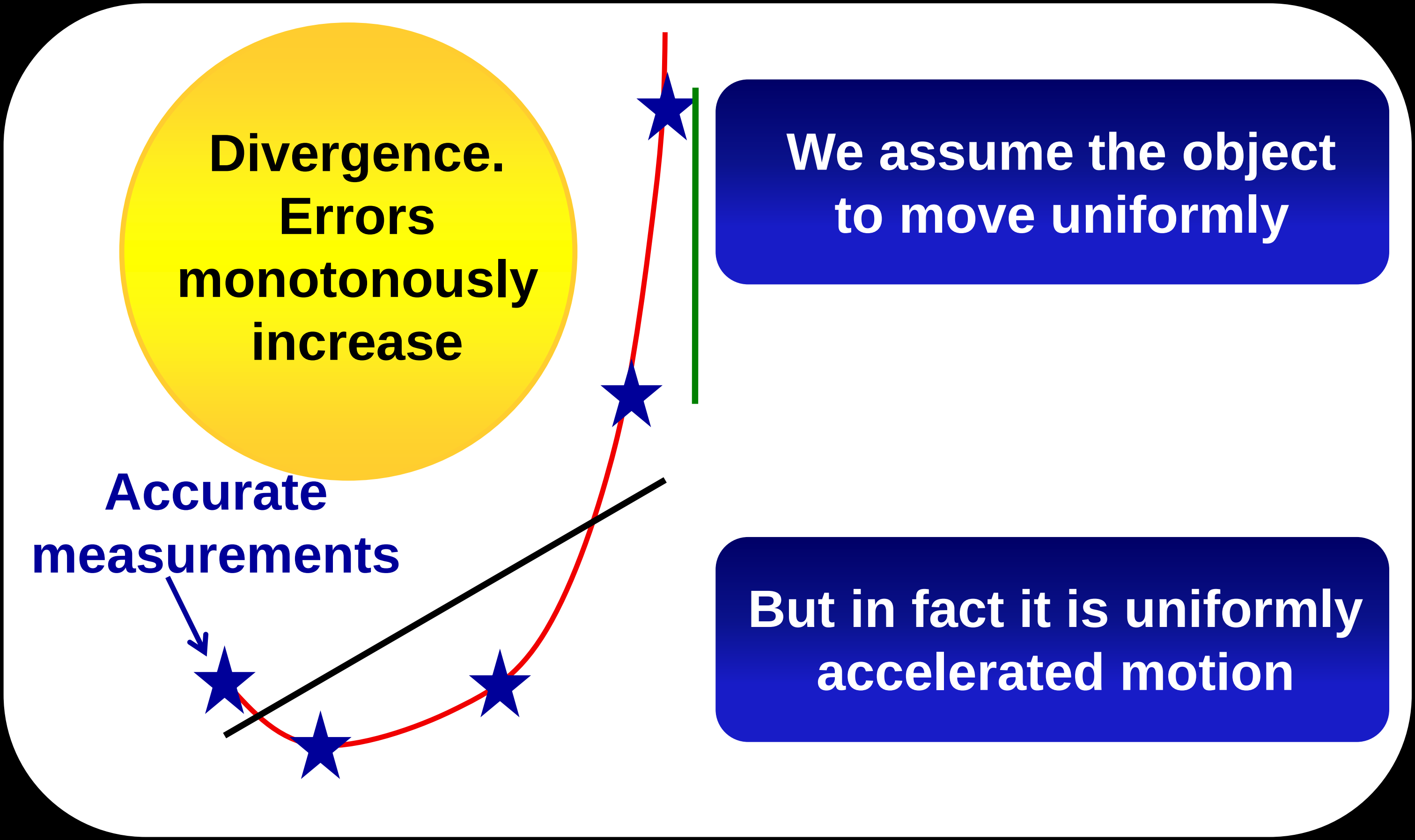
5

Sources of estimation errors



Source 1: Measurement errors → Errors of estimation are related with only measurements errors. Model of motion is accurate

Sources of estimation errors



Divergence. Errors monotonously increase

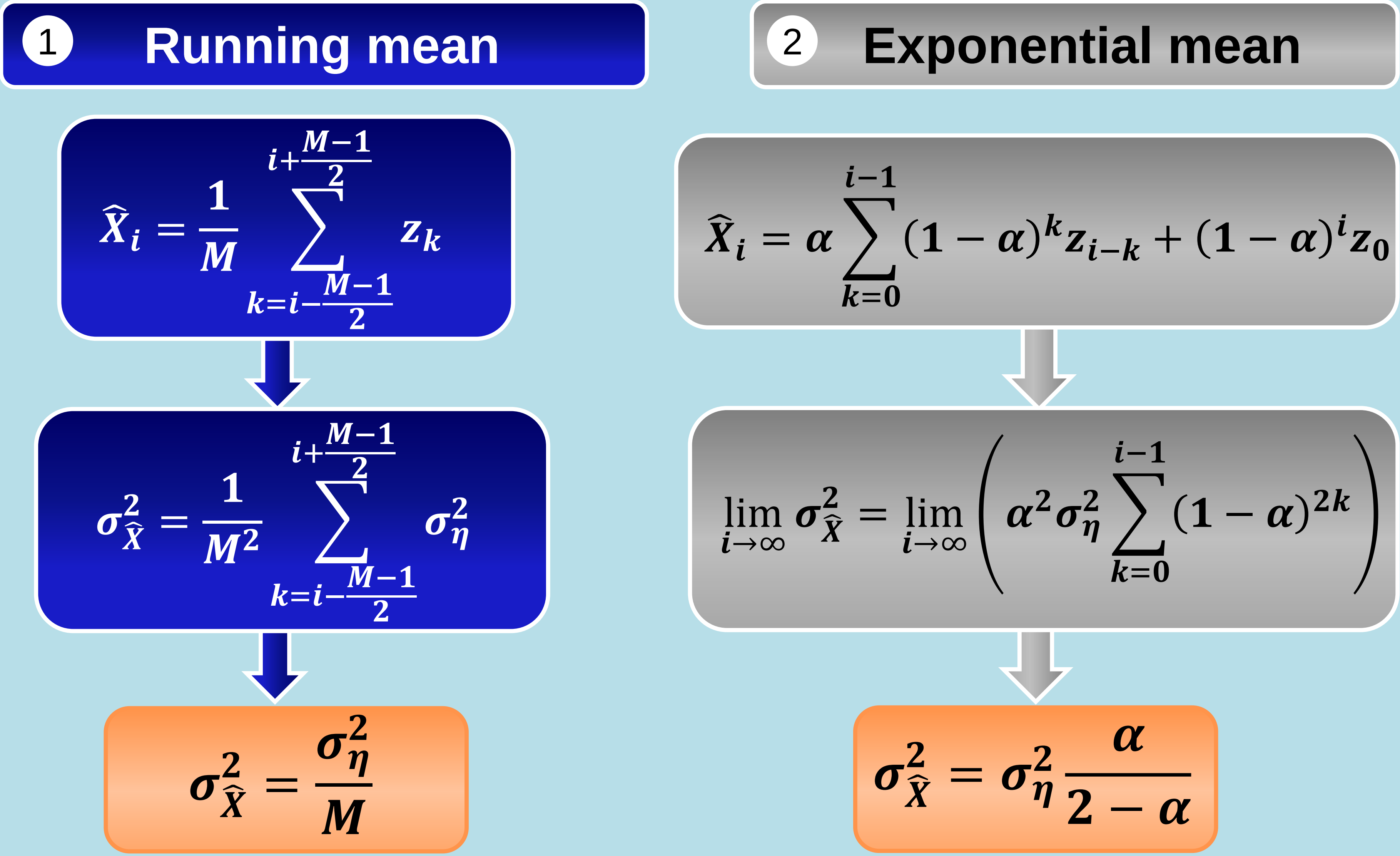
Accurate measurements

We assume the object to move uniformly

But in fact it is uniformly accelerated motion

Source 2: Methodical errors → Errors of estimation are related with errors of methods. Model of motion is inaccurate.

Source 1: Measurement errors



1 Running mean

$$\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

$$\sigma^2_{\hat{X}} = \frac{1}{M^2}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} \sigma^2_{\eta}$$

$$\sigma^2_{\hat{X}} = \frac{\sigma^2_{\eta}}{M}$$

2 Exponential mean

$$\hat{X}_i = \alpha\sum_{k=0}^{i-1}(1-\alpha)^k z_{i-k} + (1-\alpha)^i z_0$$

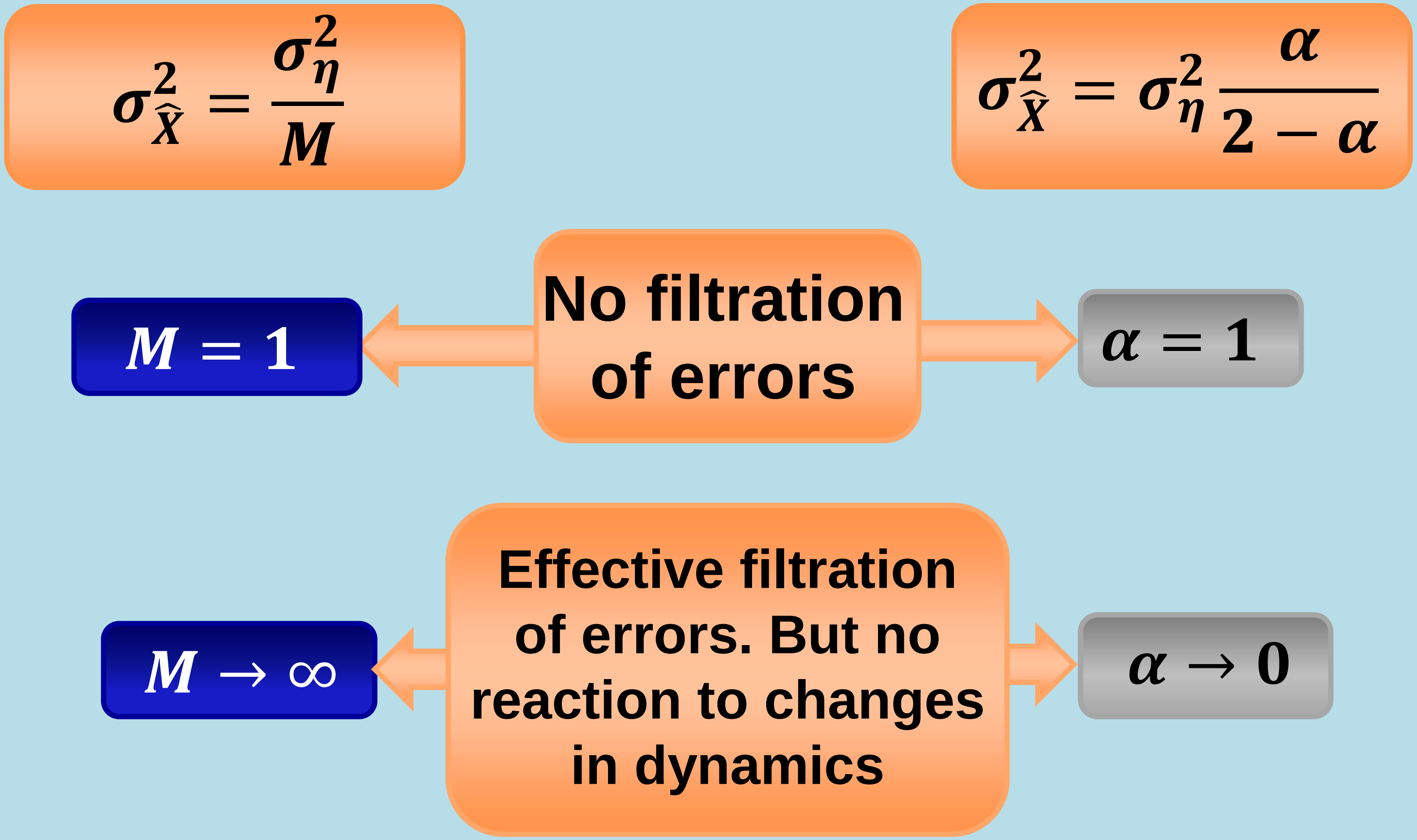
$$\lim_{i\to\infty}\sigma^2_{\hat{X}} = \lim_{i\to\infty}\left(\alpha^2\sigma^2_{\eta}\sum_{k=0}^{i-1}(1-\alpha)^{2k}\right)$$

$$\sigma^2_{\hat{X}} = \sigma^2_{\eta}\frac{\alpha}{2-\alpha}$$

Source 1: Measurement errors

1 Running mean

2 Exponential mean



$$\sigma_{\widehat{X}}^2 = \frac{\sigma_\eta^2}{M}$$

$$\sigma_{\widehat{X}}^2 = \sigma_\eta^2 \frac{\alpha}{2 - \alpha}$$

$M = 1$

No filtration of errors

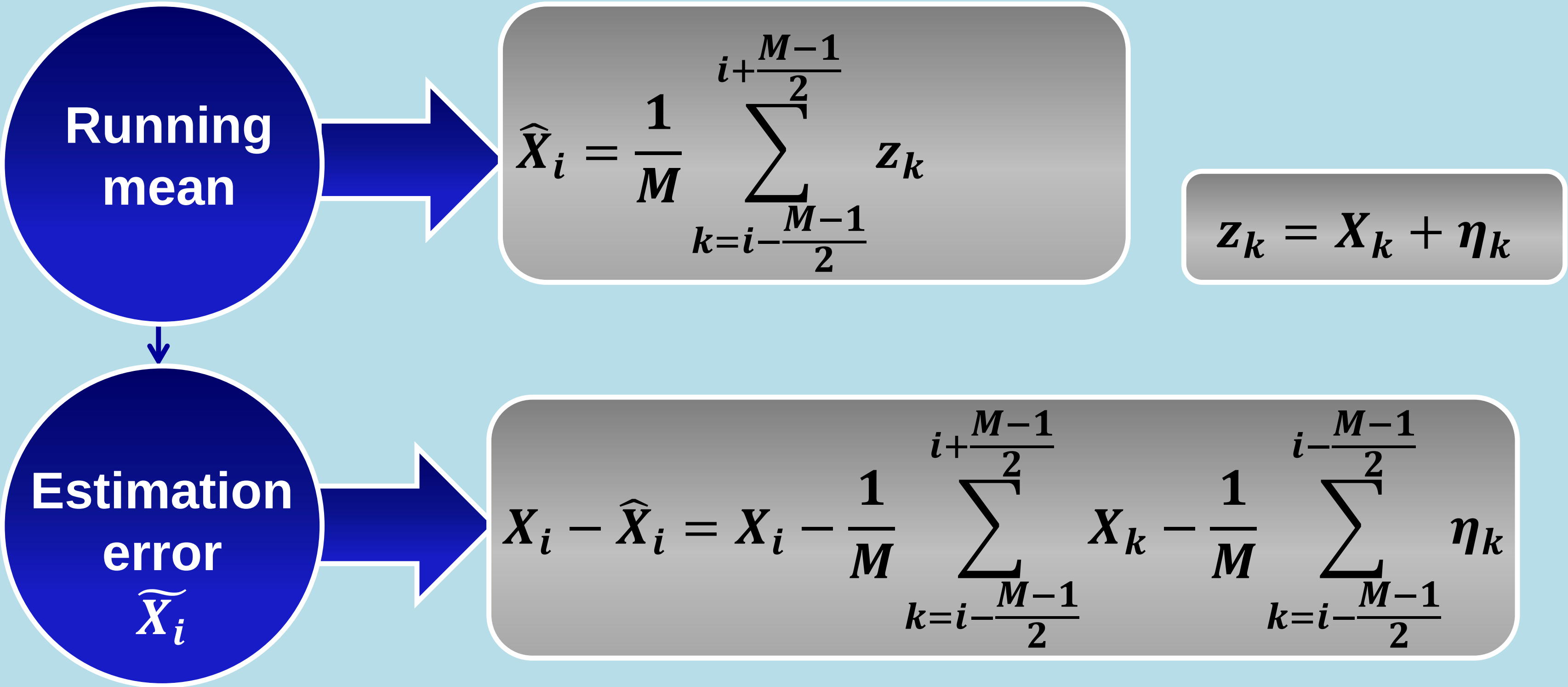
$\alpha = 1$

$M \to \infty$

Effective filtration of errors. But no reaction to changes in dynamics

$\alpha \to 0$

Source 2: Methodical errors of running mean



Running mean

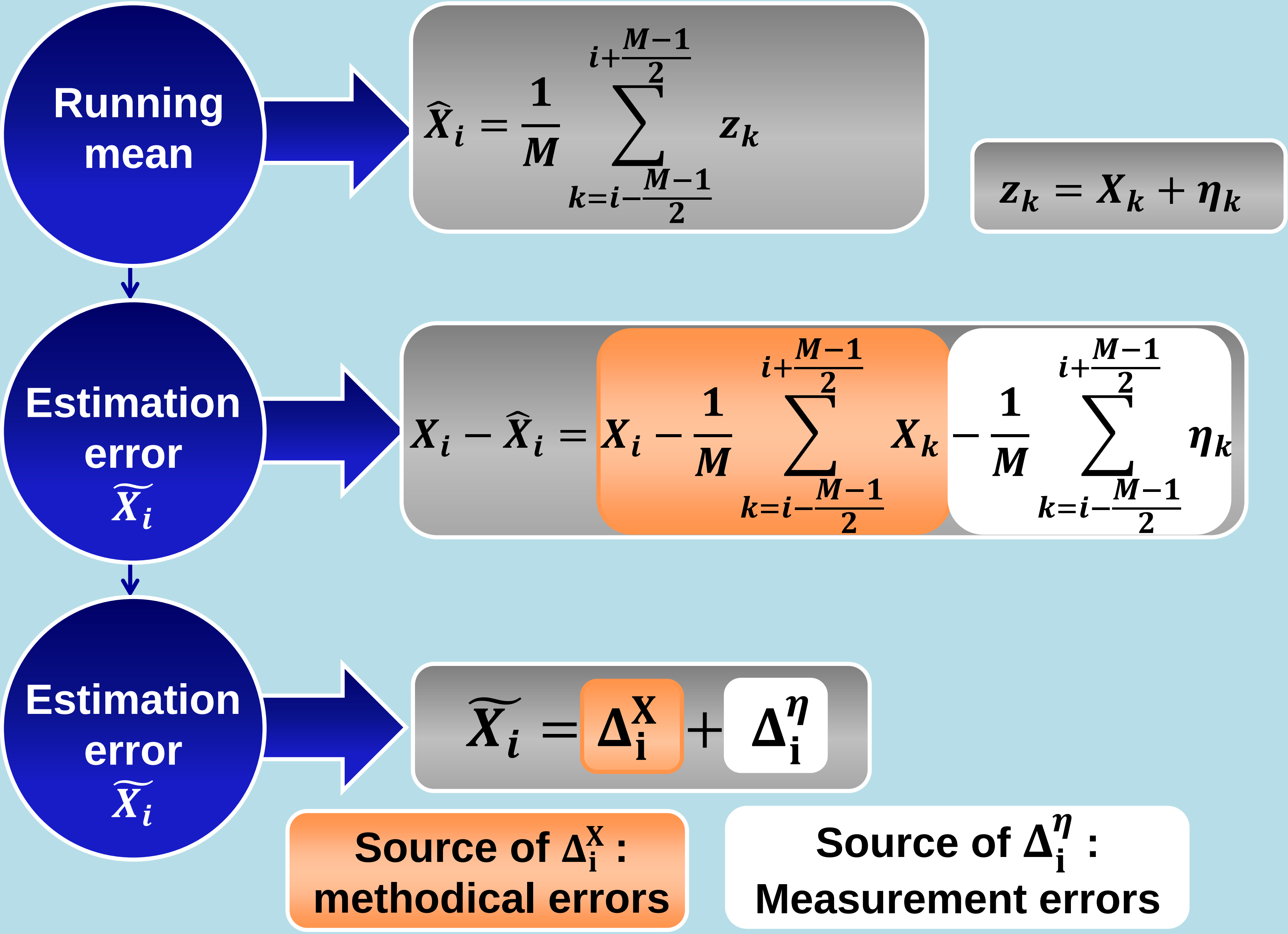
$$\widehat{X}_i = \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

$$z_k = X_k + \eta_k$$

Estimation error $\widetilde{X}_i$

$$X_i - \widehat{X}_i = X_i - \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} X_k - \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i-\frac{M-1}{2}} \eta_k$$

Source 2: Methodical errors of running mean



Running mean

$$\hat{X}_i = \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k$$

$$z_k = X_k + \eta_k$$

Estimation error $\tilde{X}_i$

$$X_i - \hat{X}_i = X_i - \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} X_k - \frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} \eta_k$$

Estimation error $\tilde{X}_i$

$$\tilde{X}_i = \Delta_{\mathrm{i}}^{\mathrm{X}} + \Delta_{\mathrm{i}}^{\eta}$$

Source of $\Delta_{\mathrm{i}}^{\mathrm{X}}$:
methodical errors

Source of $\Delta_{\mathrm{i}}^{\eta}$:
Measurement errors

Source 2: Methodical errors of running mean

$$\Delta_{\mathrm{i}}^{\mathrm{X}} = X_i - \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} X_k$$

$$9 = \sum_{k=i-4}^{i+4} 1$$

$$X_i = X_i \frac{1}{M} M = \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} X_i$$

$$\Delta_{\mathrm{i}}^{\mathrm{X}} = \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} (X_i - X_k)$$

$$X_i - \hat{X}_i = -\frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i-1} (X_k - X_i) + \frac{1}{M} \sum_{k=i+1}^{i+\frac{M-1}{2}} (X_i - X_k)$$

$k < i$

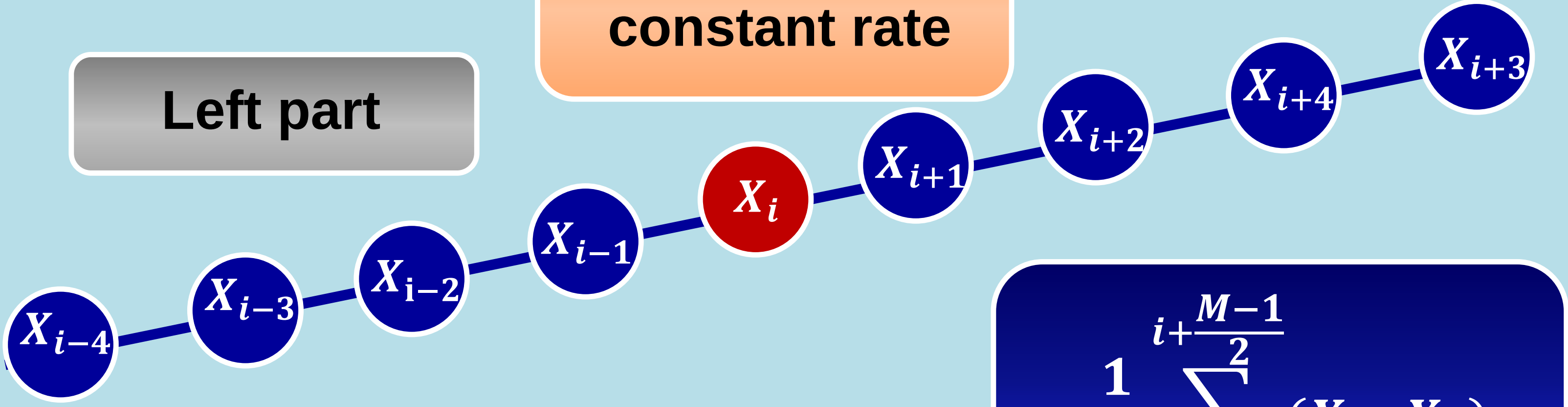
$k > i$

Source 2: Methodical errors of running mean

A process has constant rate

Left part

Right part



$$-\frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i-1}(X_k-X_i)$$

$$\frac{1}{M}\sum_{k=i+1}^{i+\frac{M-1}{2}}(X_i-X_k)$$

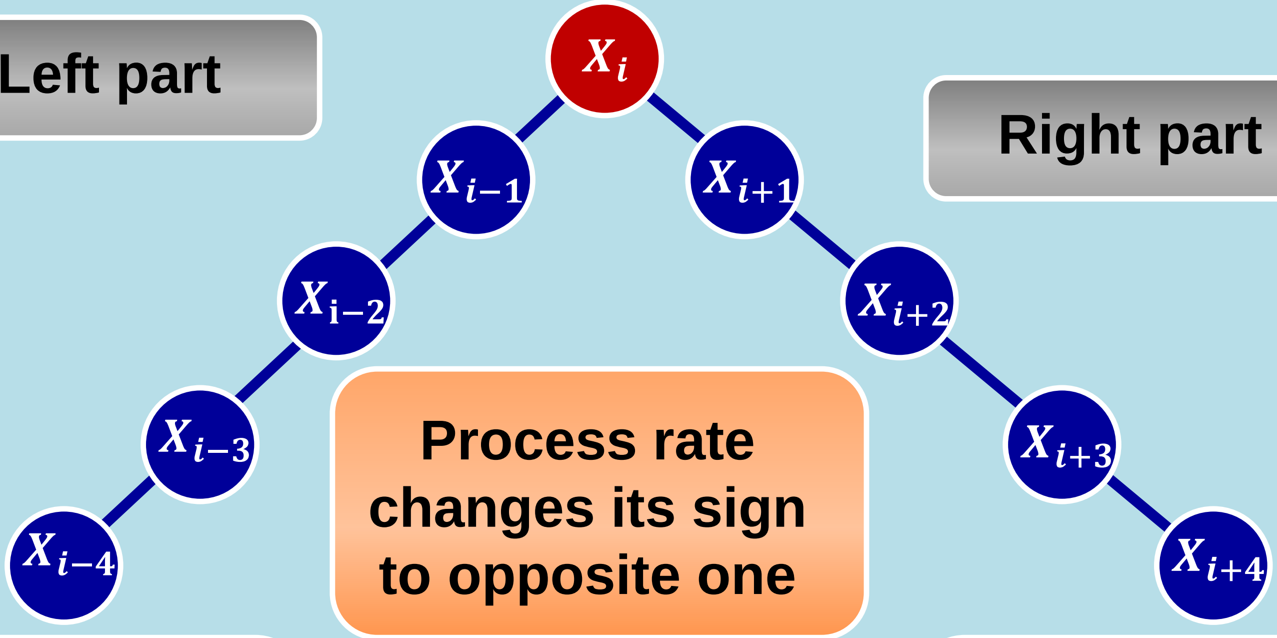
Methodical error is equal to zero

Error of right part compensates that of left part

Source 2: Methodical errors of running mean

Left part

Right part



Process rate changes its sign to opposite one

$$-\frac{1}{M}\sum_{k=i-\frac{M-1}{2}}^{i-1}(X_k-X_i)$$

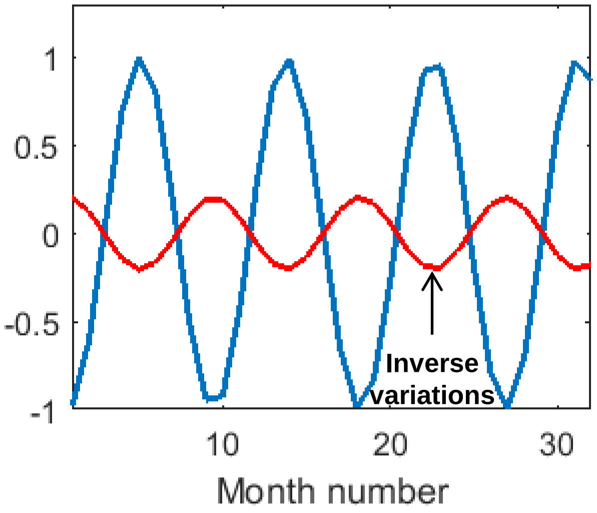
$$\frac{1}{M}\sum_{k=i+1}^{i+\frac{M-1}{2}}(X_i-X_k)$$

Methodical error is doubled → Error of right part doubles that of left part

Analysis of running mean errors

Running mean may significantly distort the dynamics of the process

9-months variation

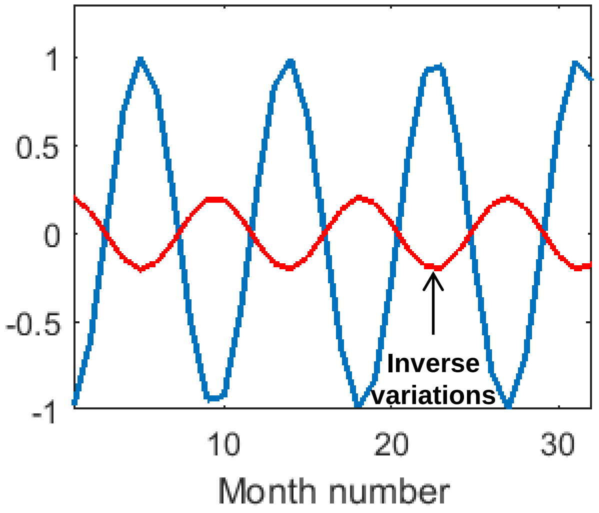


Inverse variations with periods from 6 to 12 months. Convex curve is replaced by concave curve and vice versa

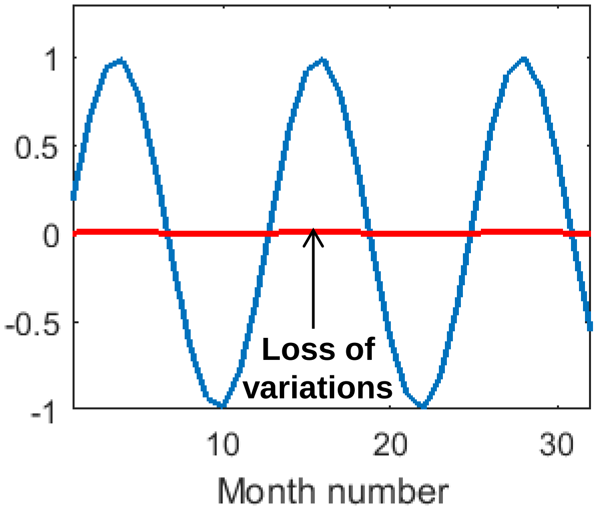
Analysis of running mean errors

Running mean may significantly distort the dynamics of the process

9-months variation



13-months variation



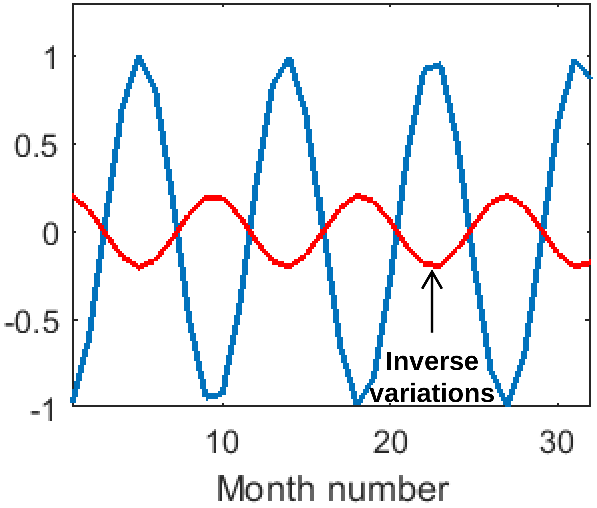
Inverse variations with periods from 6 to 12 months. Convex curve is replaced by concave curve and vice versa

Total loss of 6- and 12-month variations decreasing them to zero

Analysis of running mean errors

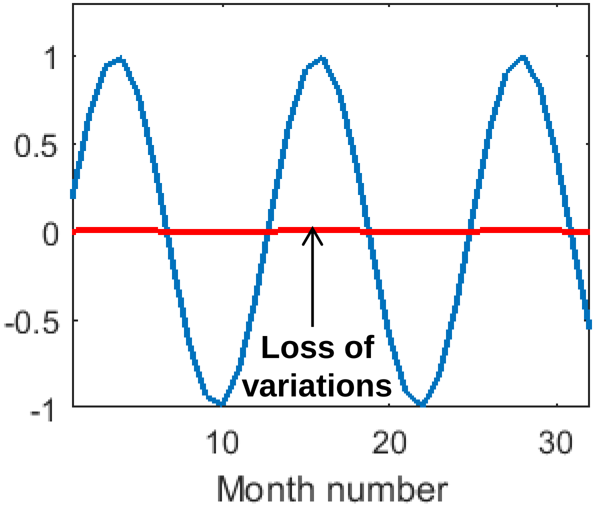
Running mean may significantly distort the dynamics of the process

9-months variation



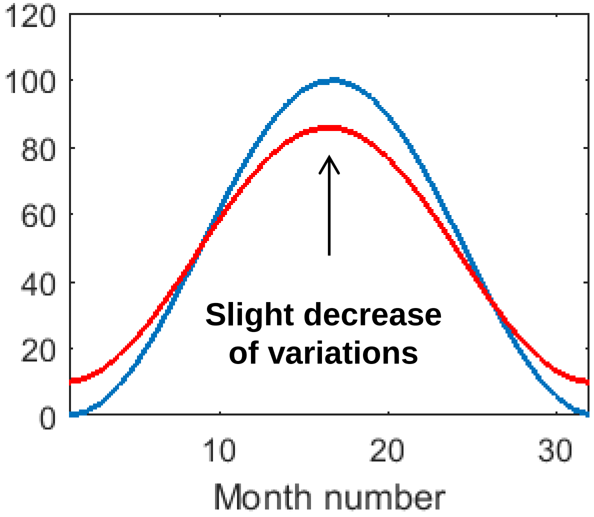
Inverse variations with periods from 6 to 12 months. Convex curve is replaced by concave curve and vice versa

13-months variation



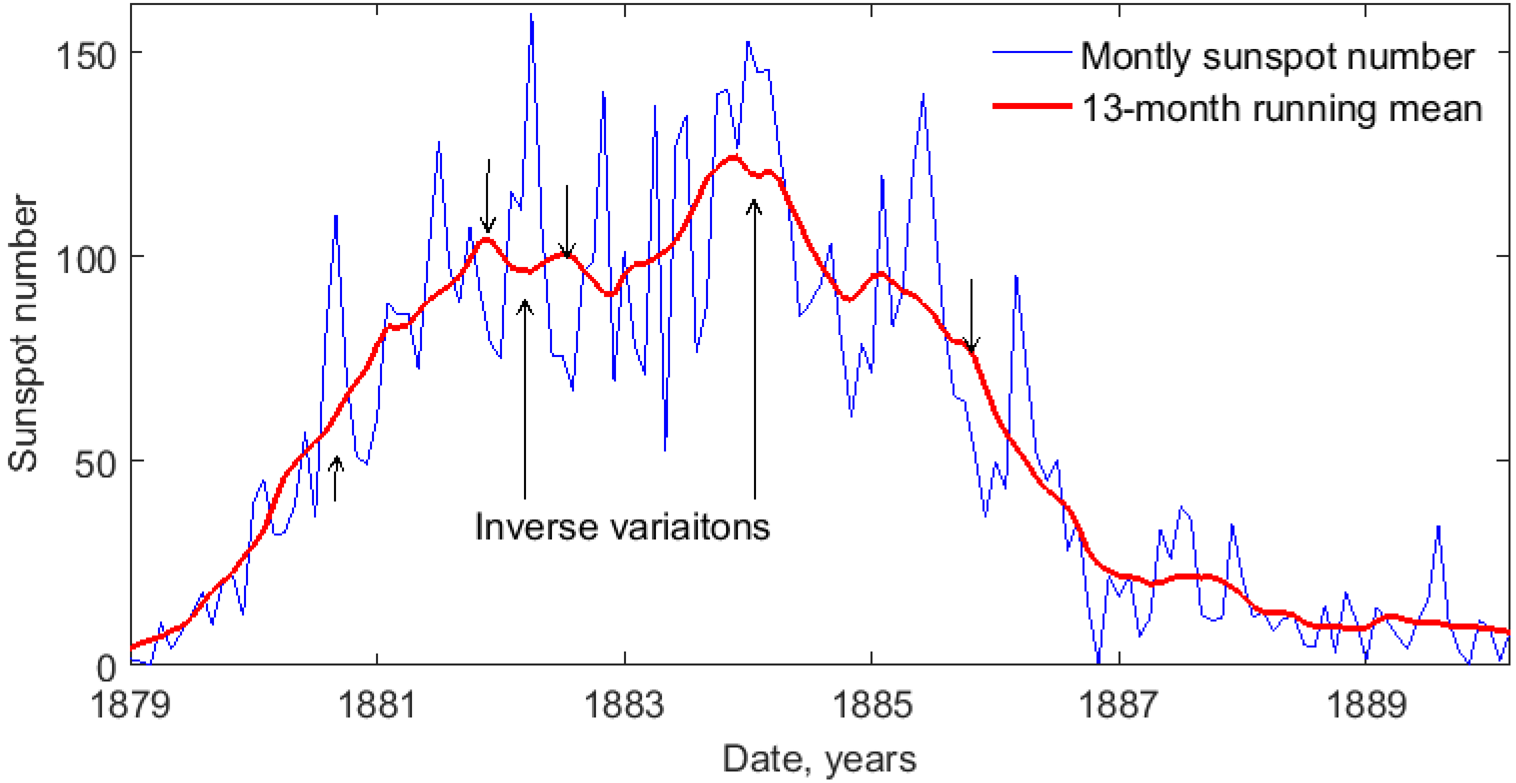
Total loss of 6- and 12-month variations decreasing them to zero

32-months variation



Period greater than running window size (13 months). The process in general is not distorted

Distortion of physics in sunspot cycle 12



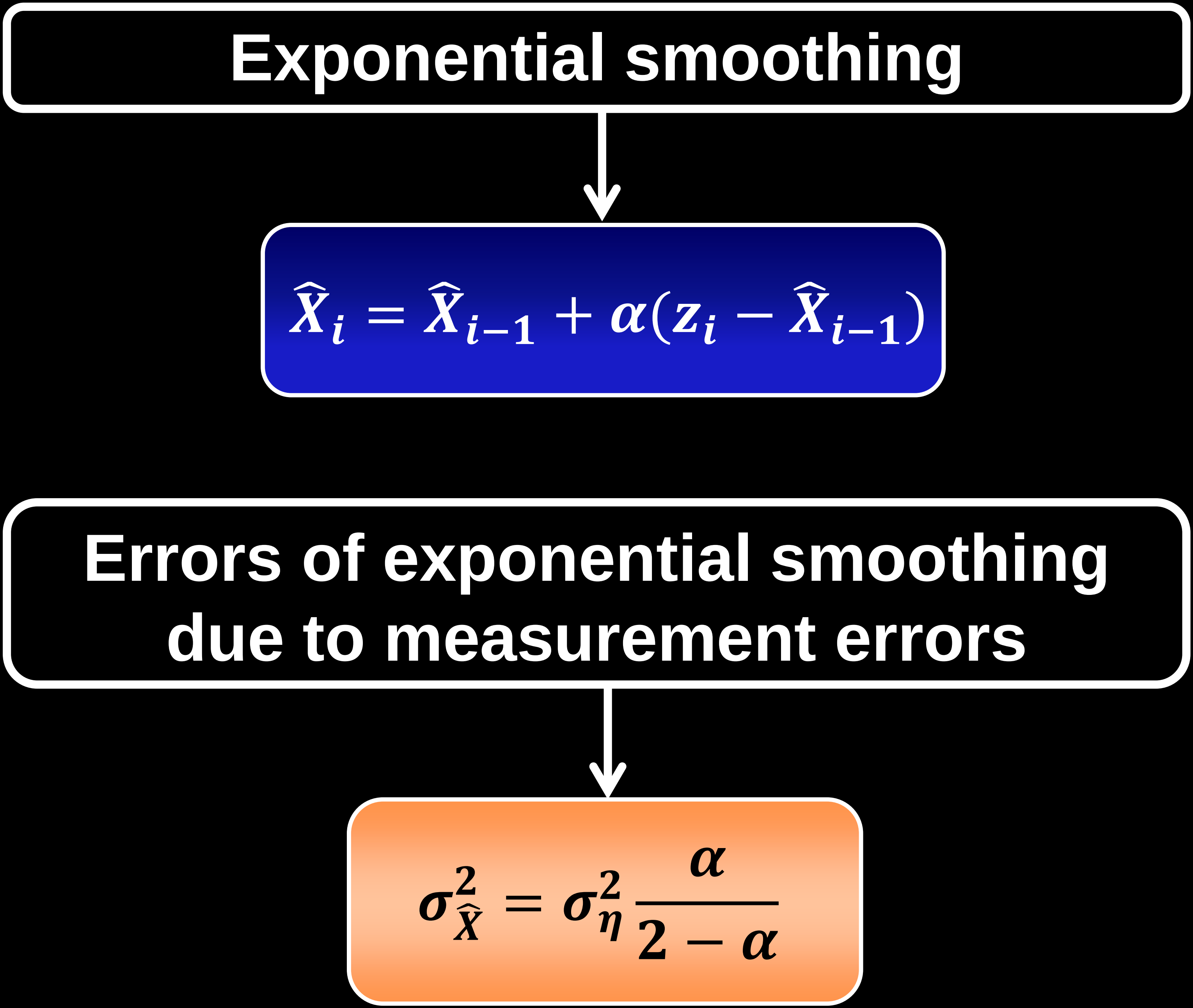
Performed analysis allows us to anticipate the errors of smoothing and getting false conclusions

Alternatives in the following topics of course

Conclusions

Don’t apply methods in blind to not fall into the trap leading to false conclusions

Even if implementation is simple, the method itself requires careful analysis



Exponential smoothing

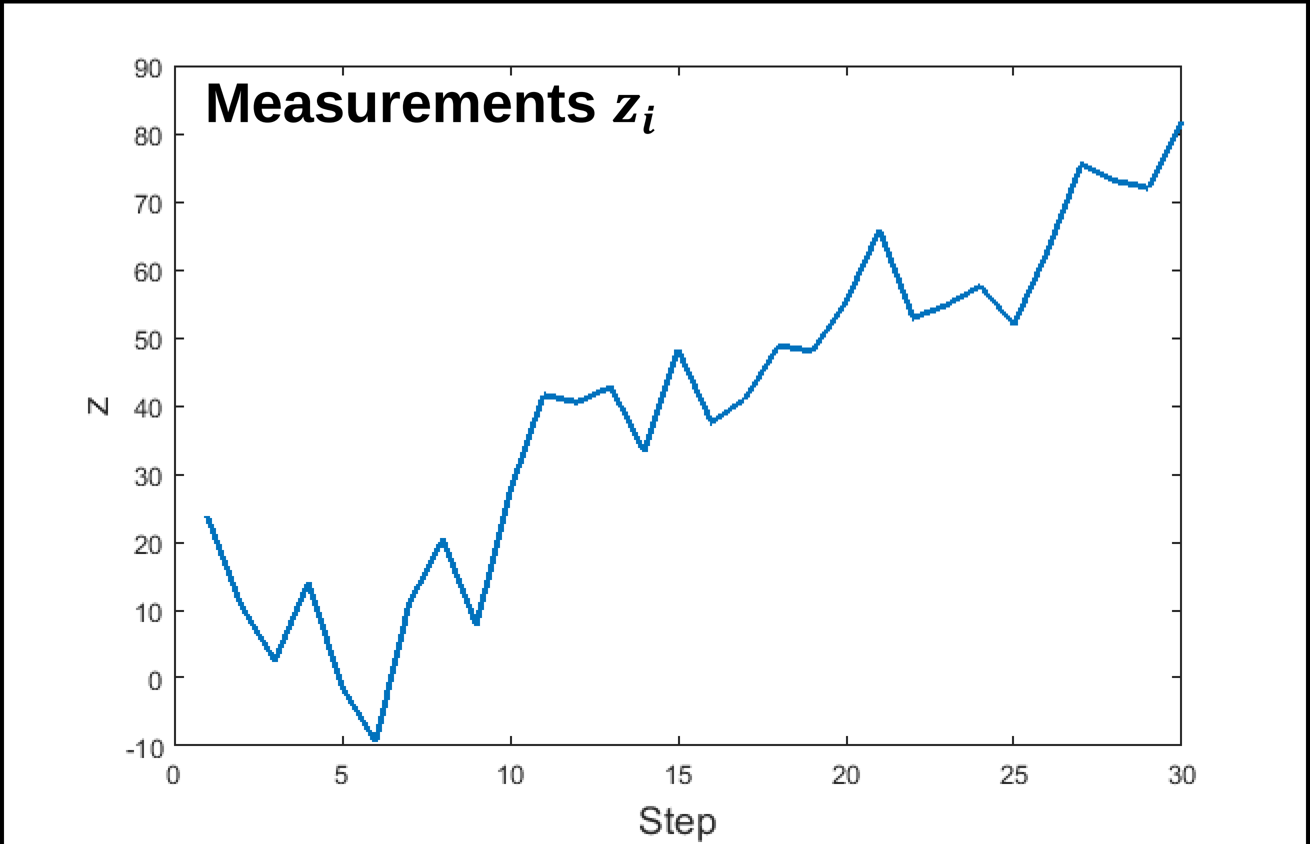
$$\widehat{X}_i = \widehat{X}_{i-1} + \alpha(z_i - \widehat{X}_{i-1})$$

Errors of exponential smoothing due to measurement errors

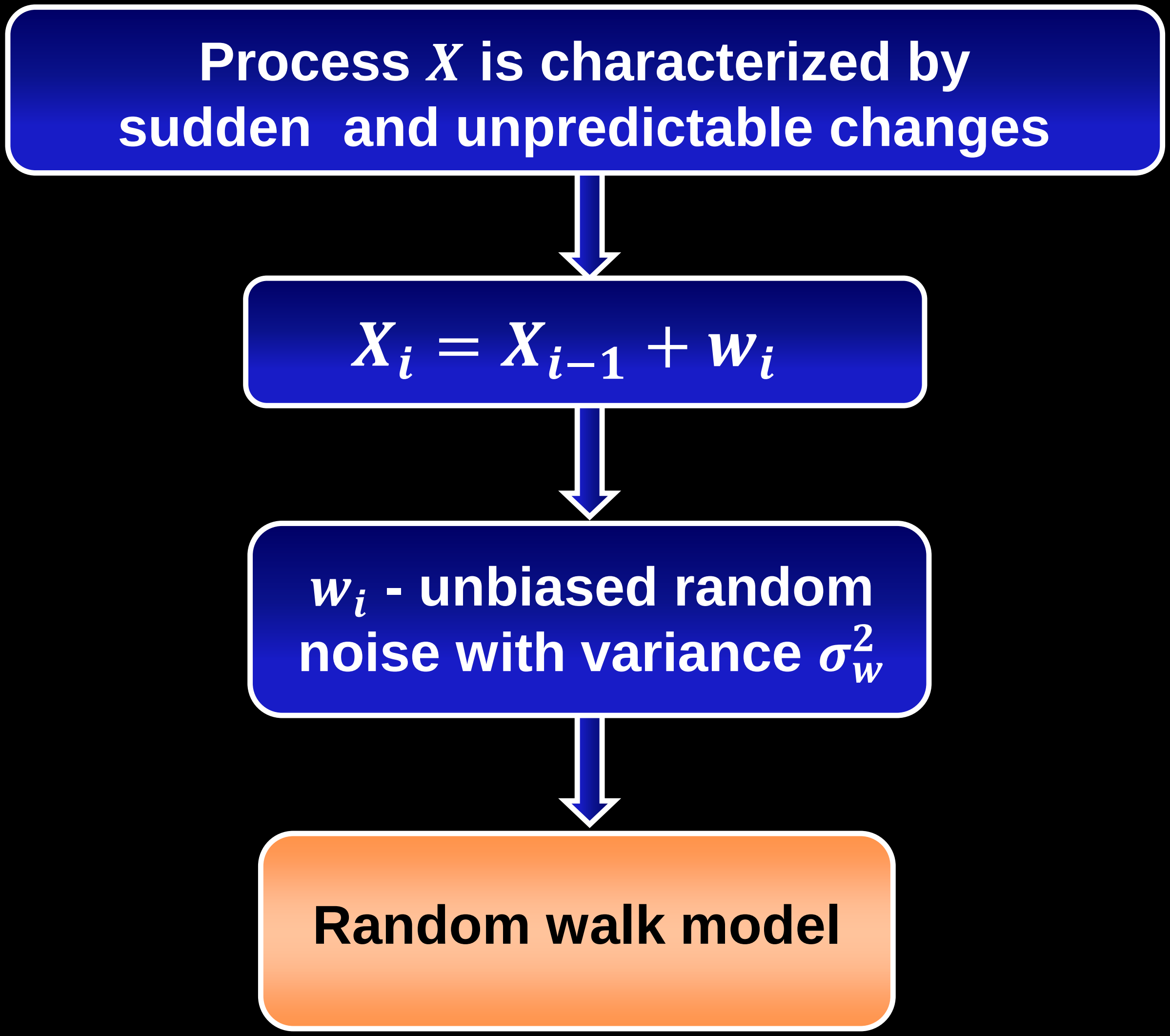
$$\sigma^2_{\widehat{X}} = \sigma^2_{\eta} \frac{\alpha}{2 - \alpha}$$

Optimal choice of smoothing constant α

Process X is characterized by sudden and unpredictable changes



Optimal choice of smoothing constant α



Process X is characterized by sudden and unpredictable changes

$$X_i = X_{i-1} + w_i$$

w_i - unbiased random noise with variance σ_w^2

Random walk model

Optimal choice of smoothing constant α

Optimal α for random walk model

$$\alpha = \frac{-\chi + \sqrt{\chi^2 + 4\chi}}{2}$$

$$\chi = \frac{\sigma_w^2}{\sigma_\eta^2}$$

σ_η^2 - variance of measurement noise

Muth J.F. (1960), Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components, J.Amer. Statist. Ass.01960.-Vol.55.-p.299.

Full error of smoothing

$$\sigma_\eta^2 a$$

$$a\sigma_\eta^2 > \sigma_\eta^2 \frac{\alpha}{2-\alpha}$$

Optimal choice of smoothing constant α

Optimal α for random walk model

$$\alpha = \frac{-\chi + \sqrt{\chi^2 + 4\chi}}{2}$$

$$\chi = \frac{\sigma_w^2}{\sigma_\eta^2}$$

σ_η^2 - variance of measurement noise

Muth J.F. (1960), Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components, J.Amer. Statist. Ass.01960.-Vol.55.-p.299.

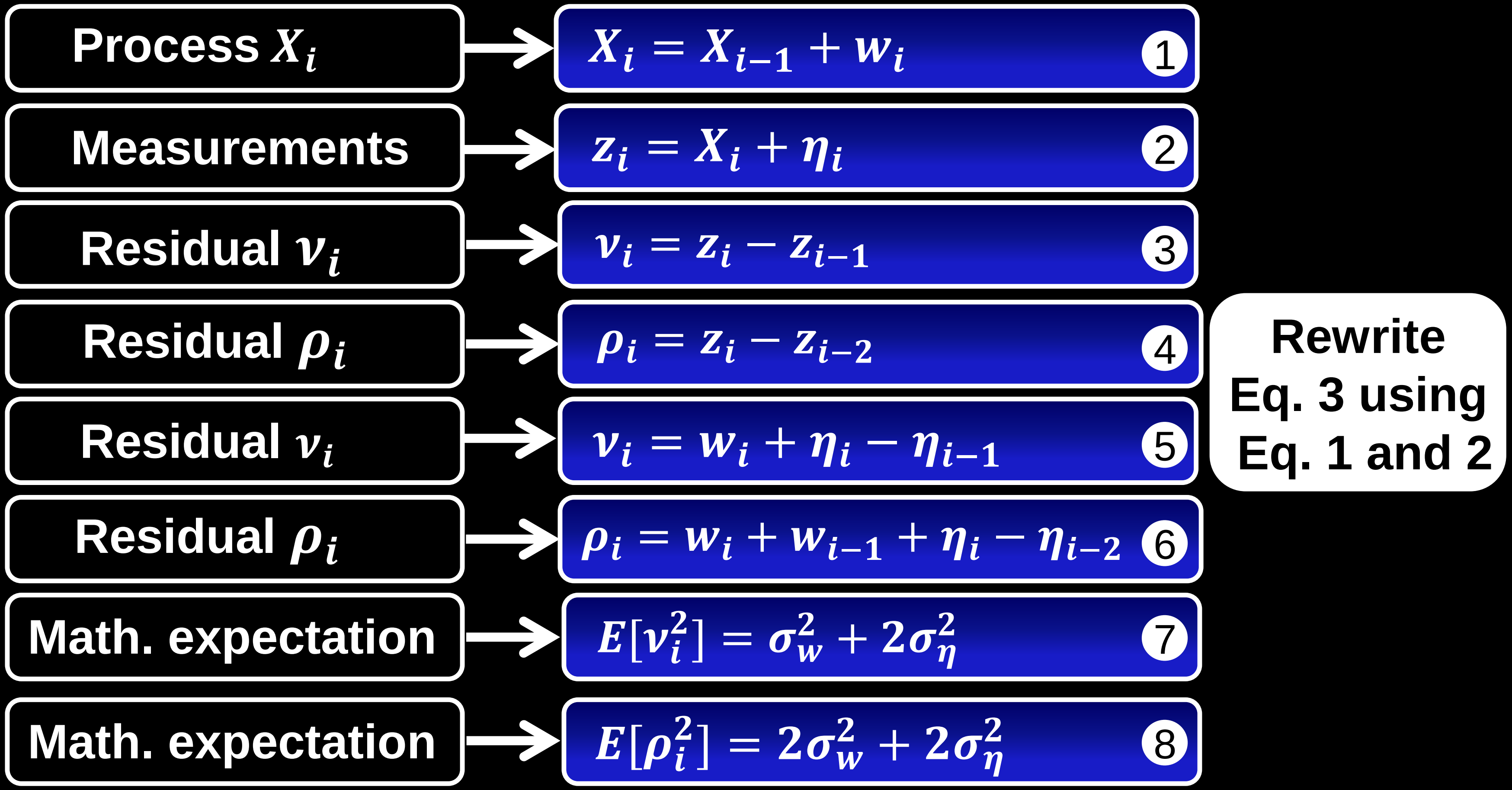
Full error of smoothing

$$\sigma_\eta^2 a$$

$$a\sigma_\eta^2 > \sigma_\eta^2 \frac{\alpha}{2-\alpha}$$

Variances σ_w^2 σ_η^2 should be identified

Identification of noise statistics σ_w^2 and σ_η^2



Process X_i → $X_i = X_{i-1} + w_i$ ①

Measurements → $z_i = X_i + \eta_i$ ②

Residual ν_i → $\nu_i = z_i - z_{i-1}$ ③

Residual ρ_i → $\rho_i = z_i - z_{i-2}$ ④

Residual ν_i → $\nu_i = w_i + \eta_i - \eta_{i-1}$ ⑤

Residual ρ_i → $\rho_i = w_i + w_{i-1} + \eta_i - \eta_{i-2}$ ⑥

Math. expectation → $E[\nu_i^2] = \sigma_w^2 + 2\sigma_\eta^2$ ⑦

Math. expectation → $E[\rho_i^2] = 2\sigma_w^2 + 2\sigma_\eta^2$ ⑧

Rewrite Eq. 3 using Eq. 1 and 2

Anderson, W. N., G. B. Kleindorfer, P. R. Kleindorfer,
and M. B. Woodroofe (1969), Consistent estimates
of the parameters of a linear system, Ann. Math. Stat., 40(3), 2064–2075.

Identification of noise statistics σ_w^2 and σ_η^2

Process X_i → $X_i = X_{i-1} + w_i$ ①

Measurements → $z_i = X_i + \eta_i$ ②

Residual ν_i → $\nu_i = z_i - z_{i-1}$ ③

Residual ρ_i → $\rho_i = z_i - z_{i-2}$ ④

Residual ν_i → $\nu_i = w_i + \eta_i - \eta_{i-1}$ ⑤

Residual ρ_i → $\rho_i = w_i + w_{i-1} + \eta_i - \eta_{i-2}$ ⑥

Rewrite Eq. 3 using Eq. 1 and 2

Math. expectation → $E[\nu_i^2] = \sigma_w^2 + 2\sigma_\eta^2$ ⑦

Math. expectation → $E[\rho_i^2] = 2\sigma_w^2 + 2\sigma_\eta^2$ ⑧

$$E[\nu_i^2] \approx \frac{1}{N-1}\sum_{k=2}^{N} \nu_k^2$$

$$E[\rho_i^2] \approx \frac{1}{N-2}\sum_{k=3}^{N} \rho_k^2$$

Consistent estimates σ_w^2 and σ_η^2 are obtained by solving system of equations (7,8)