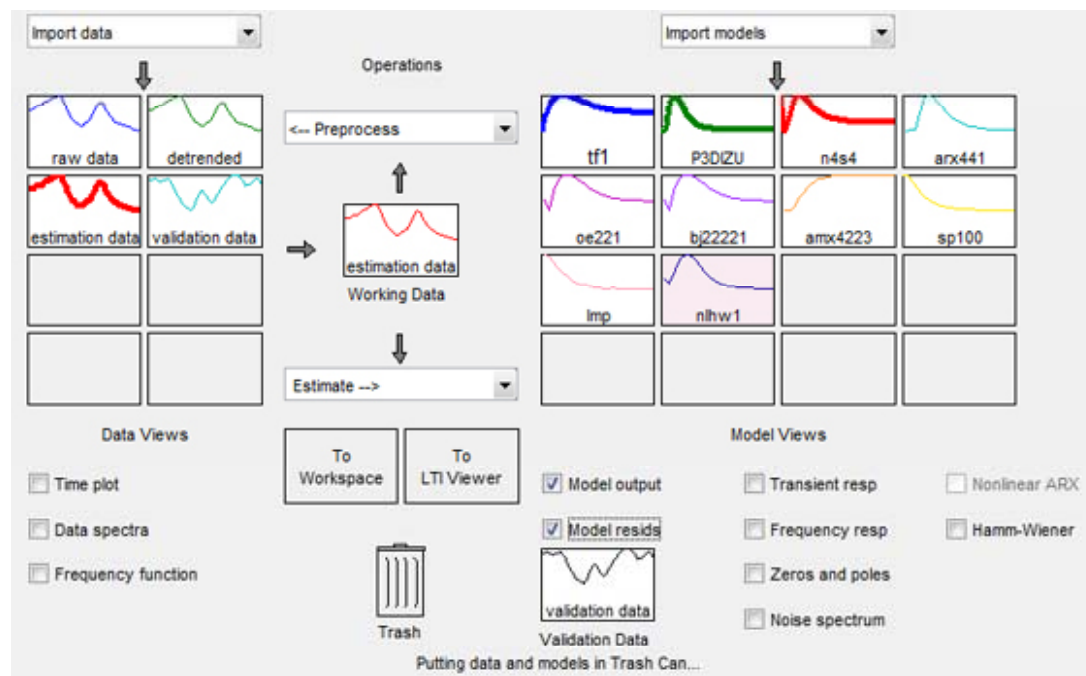
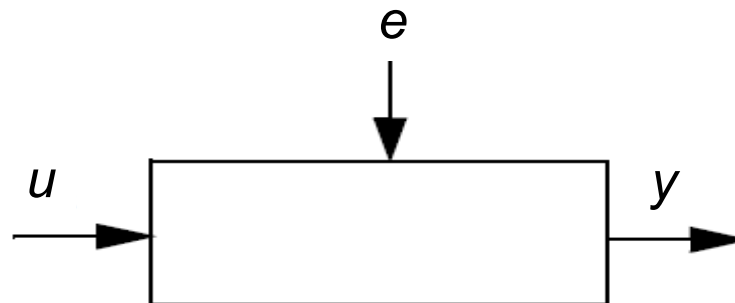


RAZVOJ DISKRETNIH MODELA IDENTIFICIRANJE MODELA



MODELI SUSTAVA

- Modeli opisuju **zavisnost izlaza o ulazima** u procesu
- Npr. na temperaturu u reaktoru utječu temperatura i protok ulaznih tokova u reaktor, protok pare u plaštu i sl.
- Na izlaz utječu i nemjerljivi signali - **poremećaji** ili **šum**.



Ulazni signalni (u), izlazni signali (y) i poremećaji (e)

VRSTE MODELA

VRSTA MODELA	PREDNOSTI	NEDOSTACI
Teorijski	▶ Dobar fizikalni uvid u vladanje procesa ▶ Primjenjivi na širokom području radnih uvjeta	▶ Potrebno dulje vrijeme za razvoj modela ▶ Problemi kod procjene modela s većim brojem parametara ▶ Nedostatak podataka o procesu
Empirijski	▶ Jednostavniji za razvoj od teorijskih modela ▶ Podaci i informacije iz realnog procesa	Problemi kod ekstrapolacije ▶ Dobar opis procesa samo na području eksperimentalnog rada
Poluempirijski	▶ Uključuju teorijsko znanje ▶ Ekstrapolacija moguća na šire područje radnih uvjeta ▶ Jednostavniji za razvoj od teorijskih modela	▶ Ne daju detaljan uvid u mehanizme procesa

DISKRETNI DINAMIČKI MODELI

- **Digitalna računala** barataju s **diskretnim** podacima, odnosno numeričkim vrijednostima funkcija danim u **vremenskim intervalima** iste duljine koji su definirani **periodom uzorkovanja** (*sampling period*);
- U računalnoj primjeni često se primjenjuju diskretni modeli kao što su **jednadžbe razlike** (*difference equations*);
- Pretvaranja kontinuiranih dinamičkih modela u diskretne provodi se primjenom **konačnih razlika** (*finite difference*);
- Razmotrimo primjer nelinearne diferencijalne jednadžbe:

$$\frac{d y(t)}{dt} = f(y, u)$$

y - izlazna varijabla
u - ulazna varijabla
t - vrijeme

DISKRETNİ DINAMIČKI MODELI

- Prethodna jednačba može se numerički integrirati uvođenjem **konačnih razlika** za **aproksimiranje derivacije**;
- Npr. aproksimacija modela prvog reda **unatražnim razlikama** (**backward difference**) glasi:

$$\frac{dy}{dt} \cong \frac{y(k) - y(k-1)}{\Delta t}$$

Δt - interval integracije (definira korisnik!)

k - diskretni korak

$y(k)$ - vrijednost $y(t)$ u $t = k \cdot \Delta t$

- Supstitucijom u jednačbu na prethodnom slajdu i proračunom $f(y, u)$ za prethodne vrijednosti y i u (tj. $y(k-1)$ i $u(k-1)$):

$$\frac{y(k) - y(k-1)}{\Delta t} \cong f(y(k-1), u(k-1))$$

$$y(k) = y(k-1) + \Delta t \cdot f(y(k-1), u(k-1))$$

DISKRETNi DINAMIČKI MODELI

- Izvedena jednađba predstavlja **jednađbu razlike prvog reda**;
- Primjenjuje se za proračun $y(k)$ na temelju vrijednosti funkcije iz prethodnog diskretnog vremenskog koraka $(k-1)$;
- Radi se o rekurzivnoj relaciji (**reccurence relation**) tj. **rekurzivnom računanju** – svaki član niza definiran je kao funkcija **prethodnih** (prošlih) **članova**;
- **Numerička integracija** provodi se uzastopnim proračunavanjem $y(k)$ za $k = 1, 2, 3, \dots$ počevši od poznatih početnih uvjeta $y(0)$ i $u(k)$.

IDENTIFICIRANJE SUSTAVA

- Parametri diskretnog modela procjenjuju se direktno iz ulaznih i izlaznih podataka dobivenih eksperimentom primjenom **linearne regresijske analize**;
- Ovaj postupak zove se **identificiranja sustava** (**system identification**);
- Primjenjuje se za modeliranje sustava koji se ne mogu opisati teorijskim modelima;
- **Primjer**: **Jednadžba razlika drugog reda**

$$y(k) = a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + b_1 \cdot u(k-1) + b_2 \cdot u(k-2)$$

služi za predviđanje $y(k)$ iz podataka dostupnih u diskretnim trenucima $(k-1) \cdot \Delta t$ i $(k-2) \cdot \Delta t$;

- Postupkom identificiranja potrebno je odrediti nepoznate parametre diskretnog modela a_1 , a_2 , b_1 i b_2 .

IDENTIFICIRANJE SUSTAVA

- Ovaj model može se izraziti u standardnom obliku:

$$y = \sum_{j=1}^p \beta_j X_j + \varepsilon$$

definiranjem:

$$\beta_1 = a_1; \quad \beta_2 = a_2; \quad \beta_3 = b_1; \quad \beta_4 = b_2$$

$$X_1 = y(k-1); \quad X_2 = y(k-2); \quad X_3 = u(k-1); \quad X_4 = u(k-2)$$

- Parametri se procjenjuju **minimiranjem** kriterija **najmanjeg kvadrata**:

$$S = \sum_{i=1}^N \left(Y_i - \sum_{j=1}^p \beta_j X_{ij} \right)^2$$

IDENTIFICIRANJE SUSTAVA

ekvivalentno, se može izraziti kao:

$$S = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

pri čemu eksponent T označava transponiranje matrice i:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ \dots \\ Y_n \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \dots \\ \beta_n \end{bmatrix}$$

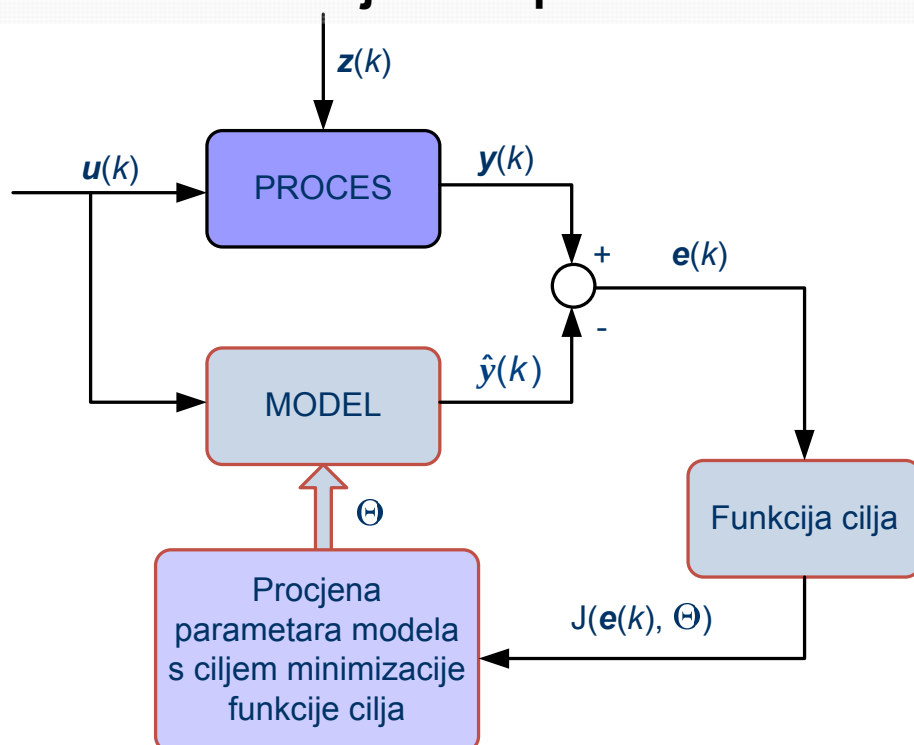
Rješenje **metodom najmanjih kvadrata** je:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

IDENTIFICIRANJE SUSTAVA

Black-box modeling

- Za izvedbu modela potrebni su eksperimentalni podaci;
- Iteracijskim postupkom optimiraju se parametri modela kojim se definira funkcijska ovisnost **izlaza modela** (simuliranih ili predviđenih) i ulaznih **izmjerenih podataka**.



$u(k)$	Ulaz u trenutku k
$y(k)$	Izlaz u trenutku k
$z(k)$	Poremećaj u trenutku k
$e(k)$	Pogreška modela u trenutku k
$\hat{y}(k)$	Modelom procijenjena vrijednost izlaza i procesa u trenutku k
$J(e(k), \theta)$	Kriterij funkcije cilja
θ	Parametri modela procesa

PRIMJER DINAMIČKOG MODELA

- Osnovna relacija je linearna diferencijalna jednačba:
npr.: $y(k) - 1,5y(k - T) + 0,7y(k - 2T) = 0,9u(k - 2T) + 0,5u(k - 3T)$

- Iz ove relacije računa se izlaz:

$$y(k) = 1,5y(k - T) - 0,7y(k - 2T) + 0,9u(k - 2T) + 0,5u(k - 3T)$$

- Izlaz u trenutku ***k*** je linearna kombinacija prošlih ulaza i izlaza;
- Identificiranjem se na temelju izmjerenih podataka računaju:
 - **koeficijenti** u jednačbi (npr. 1,5; -0,7, itd.)
 - **broj prošlih izlaza** potreban za opis (npr. dva)
 - **kašnjenje** (npr. nakon $2T$ započinje *utjecaj* u na y)
 - **broj prethodnih ulaza** (npr. dva)
- Broj prethodnih izlaza koji se uzima naziva se **red modela** (***model order***).

POSTUPAK IDENTIFICIRANJA DINAMIČKIH MODELA

- 1. Prikupljanja podataka**
- 2. Analiza i predobrada podataka**
 - grafički prikaz podataka
 - uklanjanje *outliera*, *offset*-a i linearnih trendova
 - filtriranje i generiranje podataka
 - ponovno uzorkovanje (*resampling*)
 - odabir područja koja će se analizirati
- 3. Odabir utjecajnih varijabli**
- 4. Odabir strukture modela**
- 5. Optimiranje parametara modela**
- 6. Vrednovanje modela**
- 7. Primjena modela**
 - npr. predikcija izlaza modela za napredno vođenje procesa

UKLANJENJE NEŽELJENIH KOMPONENATA

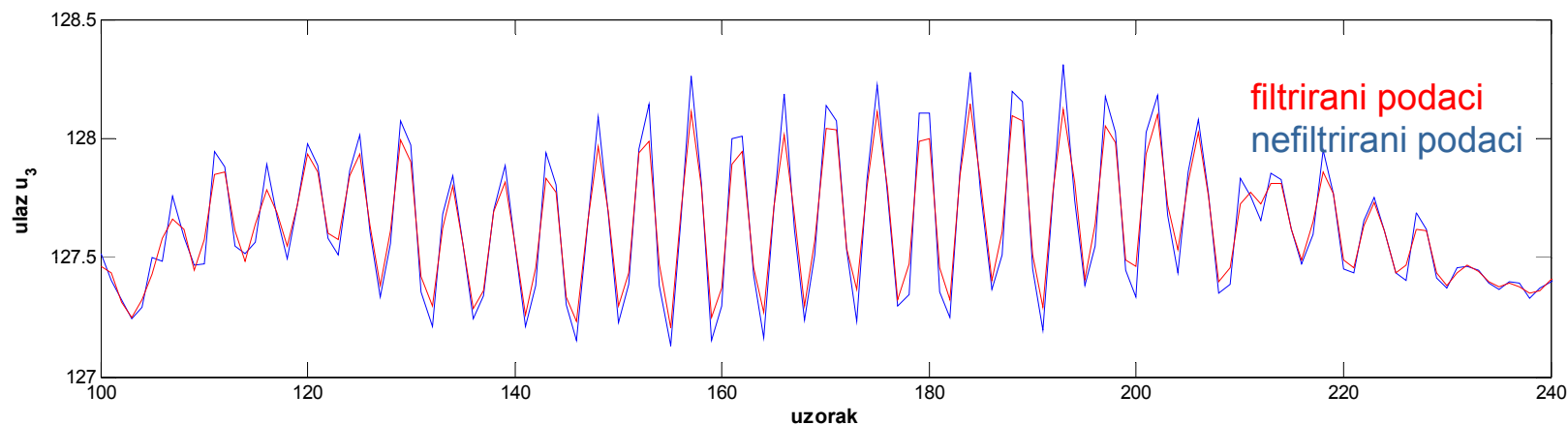
- Iz podataka se uklanjaju:
 - ekstremne vrijednosti (**outliers**),
 - **offset**,
 - linearni trendovi,
 - klizanje (**drift**) karakteristike procesa i/ili senzora.
- **Uzroci:** kvar mjerne opreme,
interferencije pri prijenosu signala,
odstupanja od radnih uvjeta procesa ,
promjene uvjeta u procesu ili sezonalne promjene,
neredovito umjeravanje mjernih pretvornika itd.
- **Metode uklanjanja:**
 - **moving window** tehnike,
 - metode najmanjih kvadrata, i sl.

FILTRIRANJE PODATAKA

Visokofrekventne smetnje uzrokuju probleme pri modeliranju.

Metode filtriranja podataka:

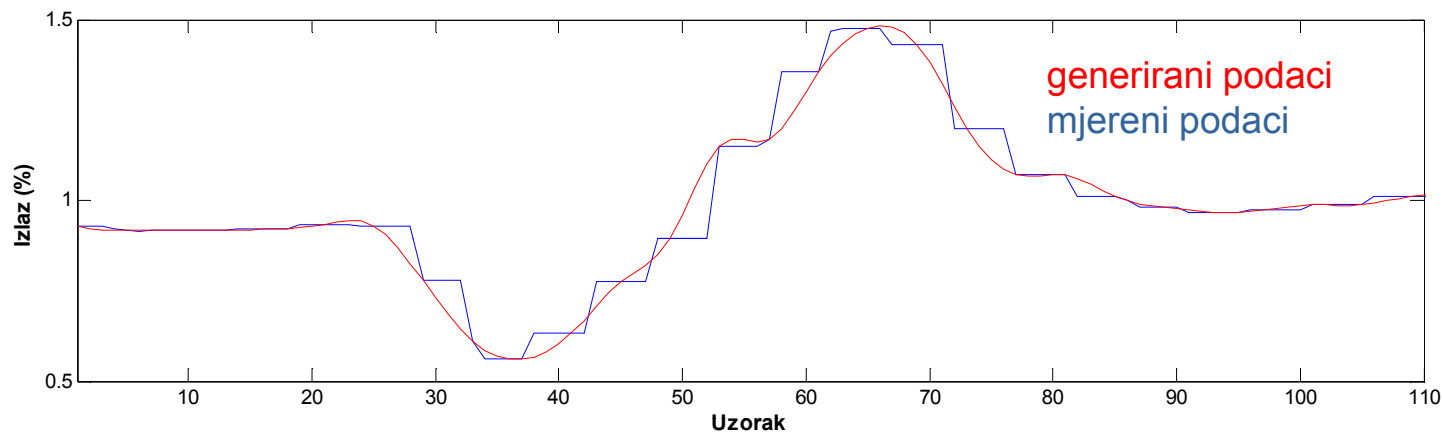
- filter prvog reda,
- Butterworth filter,
- EWMA filter (*Exponentialy weighted moving average low pass*),
- Humpel filter,
- 3-sigma filter itd.



Usporedba filtriranih i nefiltriranih podataka

NEDOSTAJUĆI PODACI

- U podacima se povremeno javljaju **nedostajući vrijednosti** (*missing values*), npr. $\pm\infty$, #####, *not a number*, *out of range*.
- **Uzroci:** kvar mjernog pretvornika, poteškoće pri prijenosu signala ili velika kompresija podataka u bazi podataka.
- **Tehnike zamjene podataka:**
 - aritmetička sredina podataka
 - viševerižinska metoda maksimalne vjerojatnosti (*maximum-likelihood*)
 - metoda Bayesian-og višestrukog unosa, itd.



Interpolirane vrijednosti dijela izlaznog signala primjenom kubnog spline-a

STRUKTURA MODELA I REGRESORA

- Odabir vrste modela ovisi prije svega o **svrsi i primjeni** modela;
- Ako je potreban približan model ili je proces blizu stacionarnim uvjetima rada, zbog jednostavnosti razvoja i lakše primjene u sustavu za vođenje najbolji izbor je **linearan model**;
- Za složene industrijske procese potrebni su **nelinearni modeli**.

STRUKTURA MODELA I REGRESORA

Regresori	Linearni modeli	Nelinearni modeli
$u(k-i)$	FIR	NFIR
$u(k-i), y(k-i)$	ARX	NARX
$u(k-i), \hat{y}(k-i)$	OE	NOE
$u(k-i), y(k-i), e(k-i)$	ARMAX	NARMAX
$u(k-i), y_s(k-i), e(k-i)$ i $e_s(k-i)$	BJ	BJ

- FIR** – Finite Impulse Response model
- ARX** – AutoRegressive model with eXogenous inputs
- OE** – Output Error model
- ARMAX** – AutoRegressive Moving Average model with eXogenous inputs
- BJ** – Box - Jenkins model

- $u(k-i)$ – prošle vrijednosti ulaza procesa
- $y(k-i)$ – prošle vrijednosti izlaza procesa
- $\hat{y}(k-i) = \hat{y}(k-i) - y(k-i)$ – prošle vrijednosti izlaza modela procesa
- $e(k-i)$ – prošle vrijednosti pogreške predikcije
- $y_s(k-i)$ – prošle vrijednosti simuliranog izlaza procesa
- $e_s(k-i) = \hat{y}_s(k-i) - y(k-i)$ – prošle vrijednosti simulirane pogreške modela

ARX model (*AutoRegressive model with eXogenous inputs*)

- linearna regresija nad prošlim mjernim uzorcima ulaznih i izlaznih signala.

$$\hat{y}(k) = [\mathbf{I} - \mathbf{A}(q)] y(k) + \mathbf{B}(q) u(k - nk)$$

$$\mathbf{A}(q) = \mathbf{I} + \mathbf{A}_1 q^{-1} + \dots + \mathbf{A}_{na} q^{-na}$$

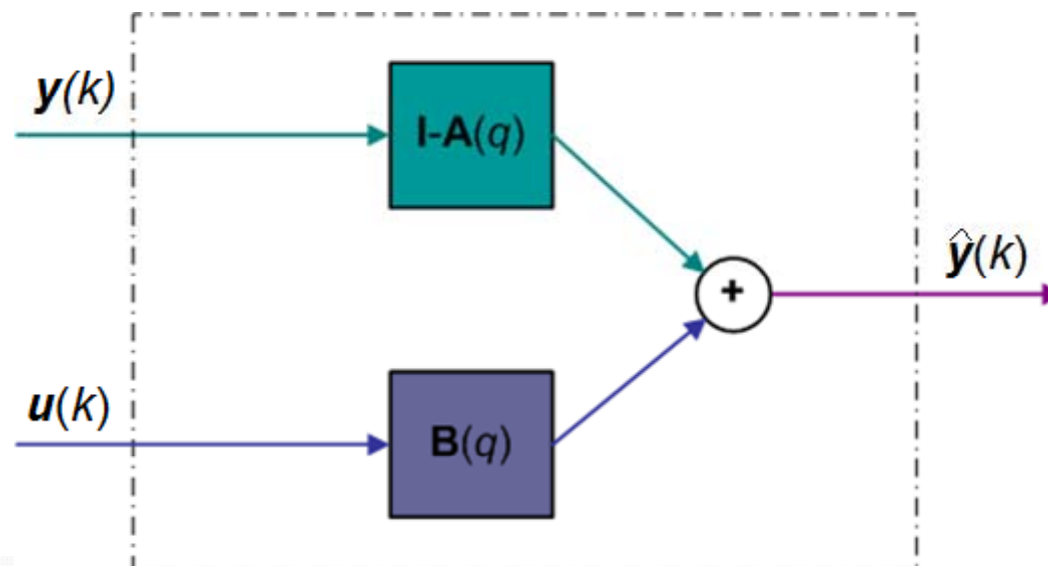
$\mathbf{A}$ je matrica koef. dimenzija: $n(y) \times n(y)$

$$\mathbf{B}(q) = \mathbf{B}_1 + \mathbf{B}_2 q^{-1} + \dots + \mathbf{B}_{nb} q^{-nb+1}$$

$\mathbf{B}$ je matrica koef. dimenzije: $n(y) \times n(u)$

q - vremenski operator
 $q^{-1}u(k) \equiv u(k-1)$

na	Broj prošlih mjerenih izlaza
nk	Zadržka ulaza
nb	Broj prošlih ulaza



OPTIMIRANJE PARAMETARA MODELA

RED MODELA I ZADRŠKA ULAZA

$(na, nb, nc, nd, nf, nk, n)$

&

KOEFICIJENTI POLINOMNOG MODELA

$(A_i, B_i, C_i, D_i, F_i)$

na	Broj prošlih mjerenih izlaza za predviđanje trenutnog izlaza (ARX, ARMAX, NARX)
nb	Broj prošlih ulaza za predviđanje trenutnog izlaza (svi modeli)
nk	Zadrška ulaza izražena brojem prošlih uzoraka (svi modeli)
nc	Broj prošlih vrijednosti signala smetnje (ARMAX, BJ).
nf	Broj prošlih vrijednosti predikcije izlaza (OE, BJ, HW)
nd	Broj prošlih vrijednosti simulirane pogreške predikcije izlaza (ARMAX, BJ).
n	Broj nelinearnih jedinica (NFIR, NARX i HW)

A_i	$A(q) = 1 + A_1q^{-1} + A_2q^{-2} + \dots + A_{na}q^{-na}$
B_i	$B(q) = B_1 + B_2q^{-1} + B_3q^{-2} + \dots + B_{nb}q^{-nb+1}$
C_i	$C(q) = 1 + C_1q^{-1} + C_2q^{-2} + \dots + C_{nc}q^{-nc}$
D_i	$D(q) = 1 + D_1q^{-1} + D_2q^{-2} + \dots + D_{nd}q^{-nd}$
F_i	$F(q) = 1 + F_1q^{-1} + F_2q^{-2} + \dots + F_{nf}q^{-nf}$

Metoda najmanjih kvadrata
Gauss-Newton
Levenberg-Marquardt
Newton-Rapson ...

VREDNOVANJE (VALIDACIJA) MODELA

Vrednovanjem modela uspoređuje se i ocjenjuje vladanje realnog procesa s procjenom modela;

Pri tome se uzimaju podaci koji **nisu primijenjeni pri razvoju modela**.

Kriteriji pri vrednovanju:

1. Srednja kvadratna pogreška (*Root mean square error*)
2. **Srednja apsolutna pogreška** (*Mean absolute error*)
3. FIT kriterij
4. **Konačna pogreška predviđanja** (*Final prediction error*)
5. Standardna devijacija (*Standard deviation*)
6. **Autokorelacijski i unakrsno-korelacijski test pogreške**
7. Frekvencijska analiza pogreške (*Histogram*)

KRITERIJI VREDNOVANJA MODELA

- Standardni kriteriji:
- Srednja kvadratna pogreška (***RMSE- Root Mean Square Error***)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

- Srednja apsolutna pogreška (***Mean Absolute Error***)

$$|\bar{e}| = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

y	izmjerene vrijednosti
$\hat{y}$	modelom procijenjena vrijednost
$\overline{y}$	srednja vrijednosti mjerene veličine
n	broj podataka

KRITERIJI VREDNOVANJA MODELA - FIT

- Kriterij slaganja modela (*FIT*)

$$FIT = \left[1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \cdot 100$$

y	Izmjerene vrijednosti
$\hat{y}$	modelom procijenjena vrijednost
$\bar{y}$	srednja vrijednosti izmjerenih vrijednosti
n	broj podataka

- Vrijednost *FIT* kriterija kreće se u rasponu od 0 do 100%.
 - 0%** → **minimalno** slaganja izmjerene vrijednosti i procjene modela.
 - 100%** → **savršeno** slaganje mjerenja i procjene modela.

KONAČNA POGREŠKA PREDVIĐANJA

- Funkcija gubitka (**V - Loss Function**)

$$V = \frac{1}{N} \sum_{k=1}^n (y(k) - \hat{y}(k, \Theta))^2$$

S povećanjem broja parametara smanjuje se vrijednost funkcije gubitka. No, to nije nužno slučaj i s pogreškom procijenjenih parametara modela.

Bolji kriterij je **konačna pogreška predviđanja**
(**FPE - Final Prediction Error**)

- modificira funkciju gubitka tako što "**kažnjava**" **kompleksnost modela** (izraženu brojem parametara modela) u odnosu na vrijednost pogreške predikcije

$$FPE = V(1 + \frac{2d}{N})$$

y	mjerenja veličina
$\hat{y}$	modelom procijenjena vrijednost
N	broj podataka
Θ	procijenjeni parametri
d	broj procijenjenih parametara

STANDARDNA DEVIJACIJA

- **Standardna devijacija** (*Standard deviation*) govori koliko su vrijednosti uzorka raspršene oko aritmetičke sredine u skupini podataka (prosječno kvadratno odstupanje od aritmetičke sredine):

$$\sigma = \sqrt{\frac{1}{n} \sum_{j=1}^n (x_i - \mu)^2}$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_i - \bar{x}_i)^2}$$

- Standardna devijacija uobičajeni je kriterij za procjenjivanje veličine slučajne pogreške i mjera ponovljivosti rezultata;
- Dobro ponovljivim rezultatima smatraju su oni čija je **s** <1% aritmetičke sredine.

AUTOKORELACIJA

- Metode identificiranja oslanjaju se na pretpostavkama o **nekoreliranosti uzoraka signala smetnje**;
- Nakon optimiranja parametara modela se pretpostavke provjeravaju **autokorelacijskim testom** (**Autocorrelation Test**) što se utvrđuje proračunom **autokorelacijske funkcije**:

$$\hat{R}_{\varepsilon\varepsilon}(\tau) = \frac{1}{N - \tau} \sum_{k=n+1}^{n+N-\tau} \varepsilon(k) \varepsilon(k + \tau)^T$$

N broj motrenih uzoraka.

$\varepsilon(k)$ odstupanje procijenjenog od izmjerenog izlaza.

- Vrijednost mora biti unutar **granica pouzdanosti** (npr. 95%) kako bi se potvrdila valjanost pretpostavke da je odstupanje modela od mjerenja nekorelirano, tj. **bijeli šum**.

UNAKRSNA KORELACIJA

- U identificiranim modelima važno je ustanoviti postoji li **unakrsna-korelacija** (**Cross-correlation test**) između pogreške modela i signala ulaza što se utvrđuje proračunom **međukorelacijske funkcije**:

$$\hat{R}_{\varepsilon u}(\tau) = \frac{1}{N - \tau} \sum_{k=n+1+\tau}^{n+N} \varepsilon(k)u(k + \tau)$$

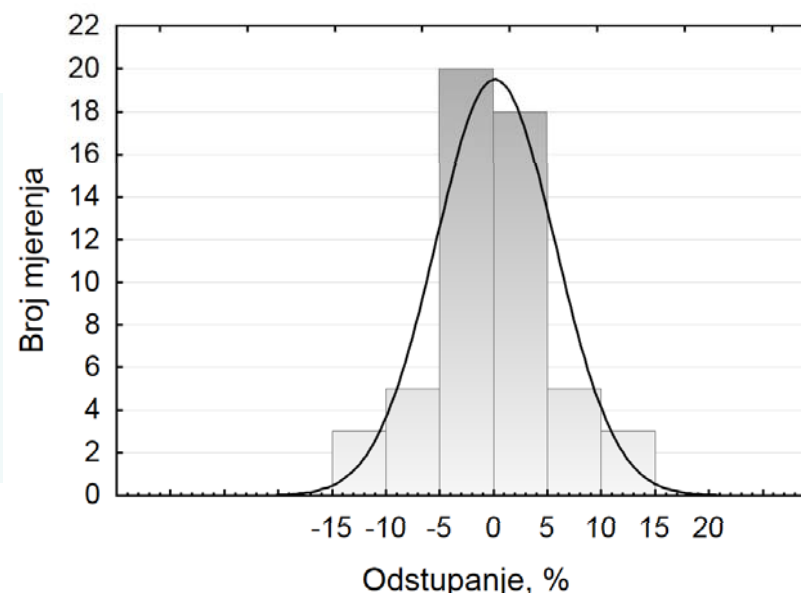
- Vrijednost mora biti u **granicama pouzdanosti** (95%) kako bi se moglo tvrditi da pogreška **modelom predviđenog izlaza** u odnosu na **stvarni izlaz ne ovisi o prošlim vrijednostima ulaza**.

HISTOGRAM

- **Histogram** prikazuje razdiobu podataka tj. učestalost pojavljivanja pojedinih vrijednost;
- Visina pravokutnika jednaka je gustoći frekvencije pojedinog intervala (razreda), odnosno frekvenciji podijeljenoj s širinom intervala;
- U slučaju normalne razdiobe podataka histogram je približno centriran oko nule i ima oblik zvona.

Što vidimo na histogramu?

- **Lokaciju podataka** – Jesu li podaci pravilno centrirani?
- **Širinu raspršenja podataka**
- **Oblik raspodjele podataka**
(simetričan, nesimetričan, dugačak rep, dva pika)
- **Ekstremno male ili velike vrijednosti**





SYSTEM IDENTIFICATION TOOLBOX - MATLAB

The screenshot displays the MATLAB 7.11.0 (R2010b) environment. The Command Window shows the following commands:

```
>> load dryer  
>> ident
```

The Workspace window shows the following data:

Name	Value	Min	Max
u2	<1000x1 double>	3.4100	6.4100
y2	<1000x1 double>	3.2008	6.2508

The System Identification Tool - Untitled window is open, showing the following options:

- Import data: [Dropdown]
- Operations: [Preprocess] [Working Data] [Estimate -->]
- Import models: [Dropdown]
- Data Views: ☐ Time plot, ☐ Data spectra, ☐ Frequency function
- Model Views: ☐ Model output, ☐ Model residuals, ☐ Transient resp, ☐ Frequency resp, ☐ Zeros and poles, ☐ Noise spectrum, ☐ Nonlinear ARX, ☐ Hamm-Wiener
- Validation Data: [Click acknowledged. No action invoked.]