

## Laskuharjoitus 9, tehtävä 6

Tämä ohje sisältää vaihtoehtoisen tavan laskuharjoituksen 9 tehtävän 6 mallien estimoimiseen käyttäen MATLAB:in System identification toolboxin System Identification Tool:ia. Myös tehtävän ratkaisupaperista löytyvä ohje kannattaa lukea.

Ladataan ensin tehtävän 6 aineisto MATLAB:in työskentelytilaan (workspace) komennolla

```
load las09t7data.mat
```

Tällä komennolla saat työskentelytilaan vektorit  $\mathbf{u}$  ja  $\mathbf{y}$ .

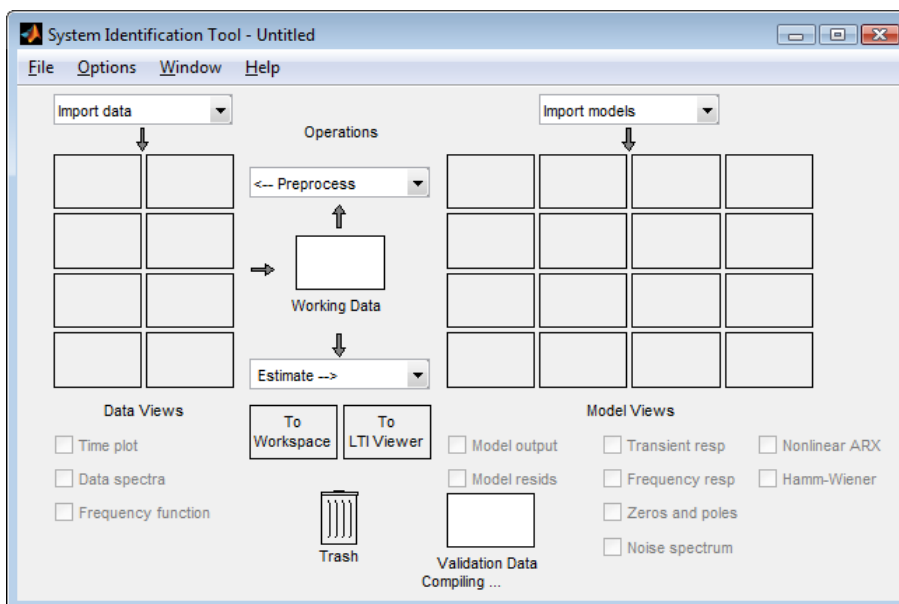
Käynnistetään nyt MATLAB:in System identification tool komennolla

```
ident
```

Näkyviin tulee työkalun päävalikko, joka on esitetty Kuvassa 1. Lisää tietoa työkalun käytöstä saat komennolla

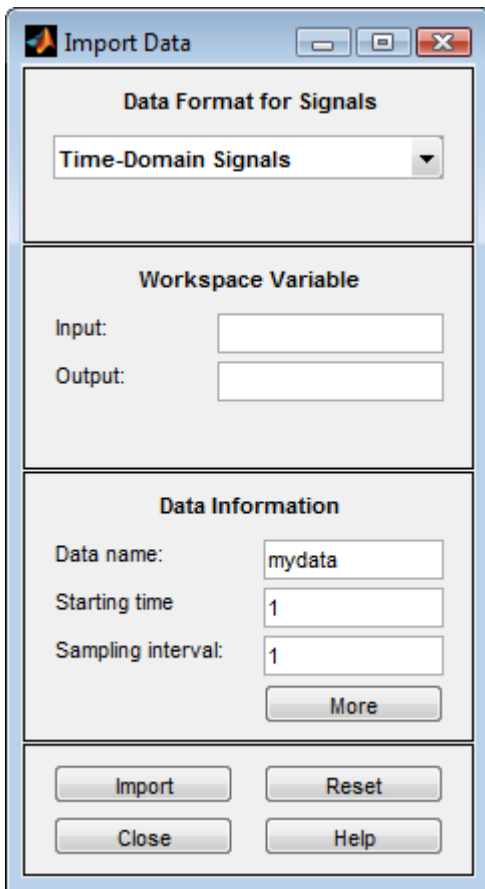
```
doc ident
```

Esimerkiksi ohje, joka löytyy päävalikosta kohdasta "Getting Started" -> "Tutorial – Identifying Linear Models Using the GUI" voi olla hyödyllinen.



Kuva 1: System identification tool:in päävalikko.

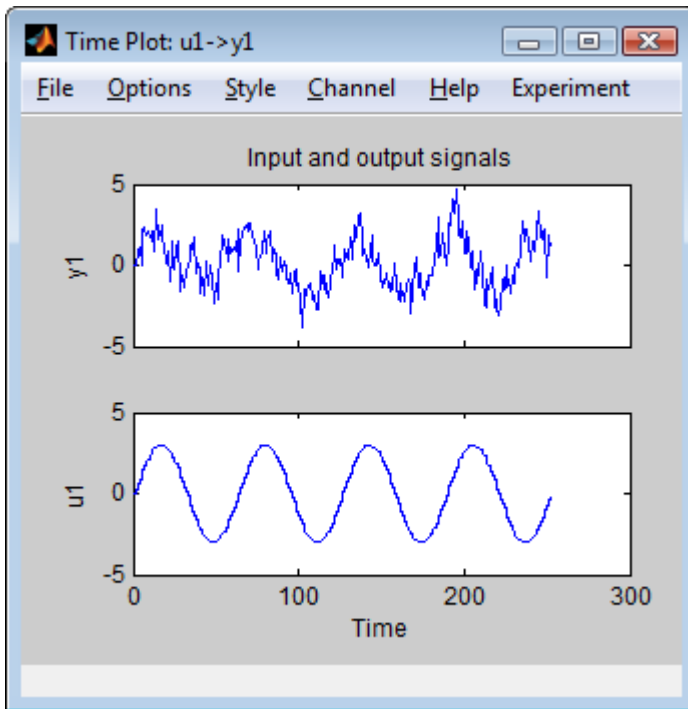
Tuodaan nyt työkaluun tehtävän aineisto. Valitaan vasemmalta "Import data"-valikosta "Time domain data...", jonka jälkeen esillä on Kuvassa 2 esitetty valikko.



Kuva 2: Aikasarjan tuominen System identification tool:iin.

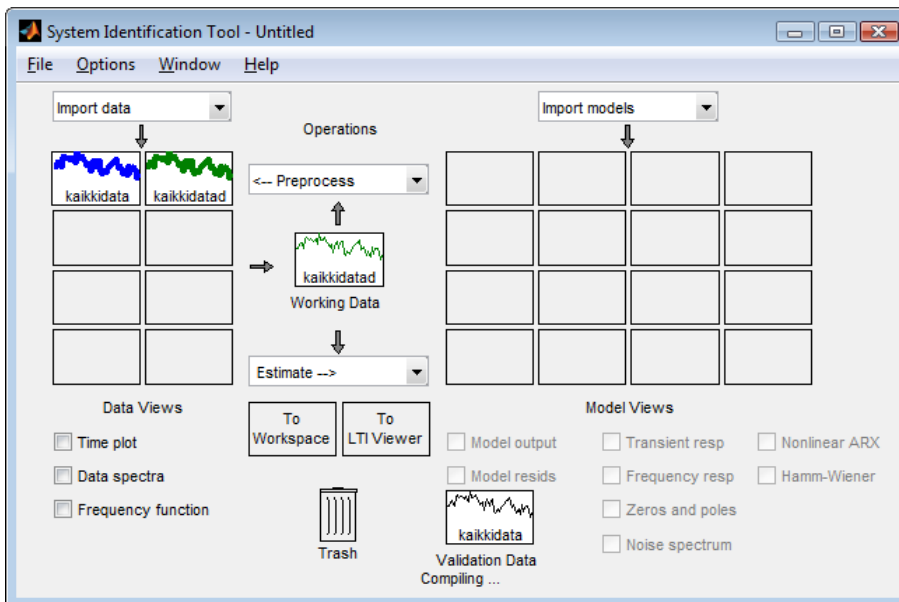
Kirjoitetaan nyt "Input"-laatikkoon **u**, ja "Output"-laatikkoon **y**. Nämä muuttujat vastaavat aikaisemmin MATLAB:in työskentelytilaan tuomiemme vektoreiden nimiä. Annetaan aineiston nimeksi "kaikkidata" . Aloitus aika (starting time) ja näytteistysväli (sampling interval) voivat tässä tapauksessa olla molemmat 1. Tuodaan aineisto työkaluun painamalla nappulaa "Import". Huomaa, että aineisto on myös tullut näkyviin "Working data" ja "Validation data"-laatikoihin.

Piirretään kuvaaja inputista ja outputista rastittamalla ruutu "Time plot" aineistoalueen alapuolelta. Näkyviin tulee kuvaaja, jossa on esitetty aikasarja aineiston input- ja output-muuttujan arvoista. Tämä on esitetty Kuvassa 3.



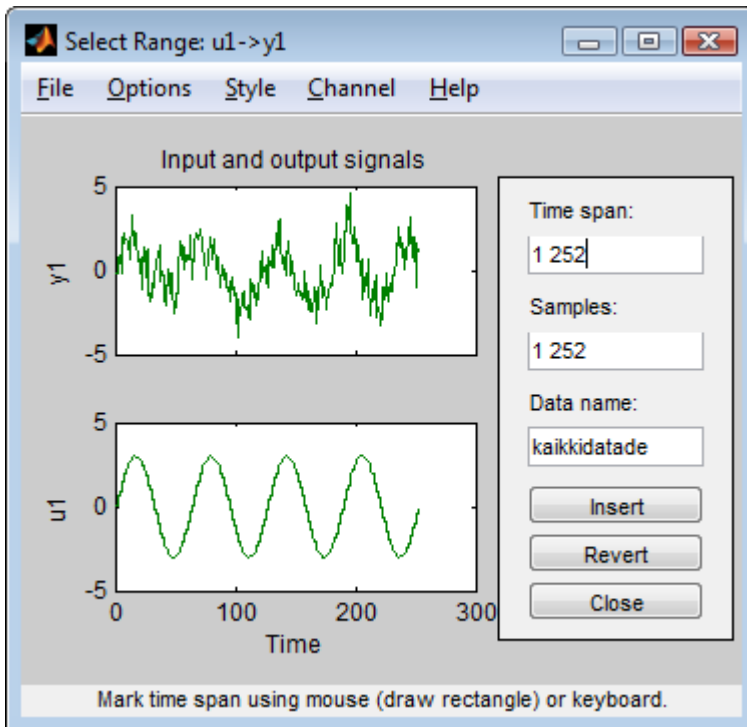
Kuva 3: Aikasarja tehtävän input- ja outputmuuttujan arvoista.

Seuraavaksi esikäsitellään aineisto, eli vähennetään siitä keskiarvo ja jaetaan se estimointi- ja validointiaineistoon. Koska aineistomme on "Working data"-ikkunassa, voimme valita keskeltä "Preprocess"-valikosta vaihtoehdon "Remove means". Tämä tuottaa aineiston "kaikkidatad". Siirretään nyt aineisto "kaikkidatad" "Working data"-laatikkoon klikkaamalla ja raahaamalla. Tämän jälkeen työkalun päävalikon pitäisi olla Kuvan 4 kaltainen.



Kuva 4: Työkalun ikkuna, kun nolakeskiarvoistettu aineisto "Working data"-laatikossa.

Jaetaan aineisto estimointi- ja validointiaineistoihin. Valitaan nyt keskeltä "Preprocess"-valikosta "Select range". Operaatio kohdistuu nyt nolakeskiarvoistettuun aineistoon "kaikkidatad" koska olemme siirtäneet sen "Working data"-laatikkoon. "Select range"-ikkuna on esitetty Kuvassa 5.



Kuva 5: Select range-ikkuna.

Select range-ikkuna piirtää valitun aineiston input- ja outputmuuttujan ajan funktiona. Koska aikaisemmin valitsimme näytteistysväliksi arvon 1, sisältävät "Time span" ja "Samples"-laatikot saman informaation, eli ne kertovat aikasarjan aloitus- ja lopetusajankohdan.

Kirjoitetaan "Data name"-laatikkoon nimi "estimointi" ja "Samples"-laatikkoon

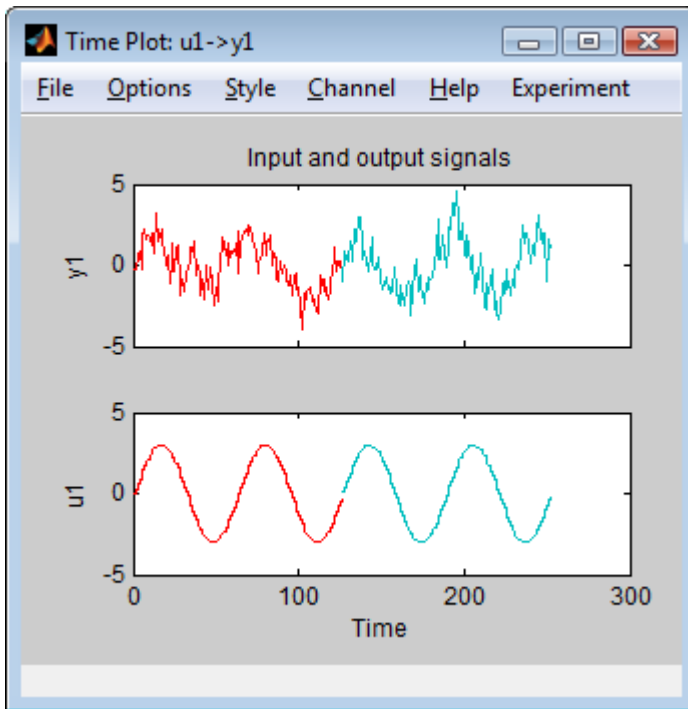
1 126

Tämän jälkeen painetaan nappulaa "Insert". Tämä luo System identification tool:in pääikkunaan uuden aineiston, joka koostuu koko aineiston ensimmäisestä puolikkaasta. "Select range"-ikkuna jää vielä auki, joten kirjoitetaan "Data name"-laatikkoon nyt "validointi", ja "Samples"-laatikkoon

127 252

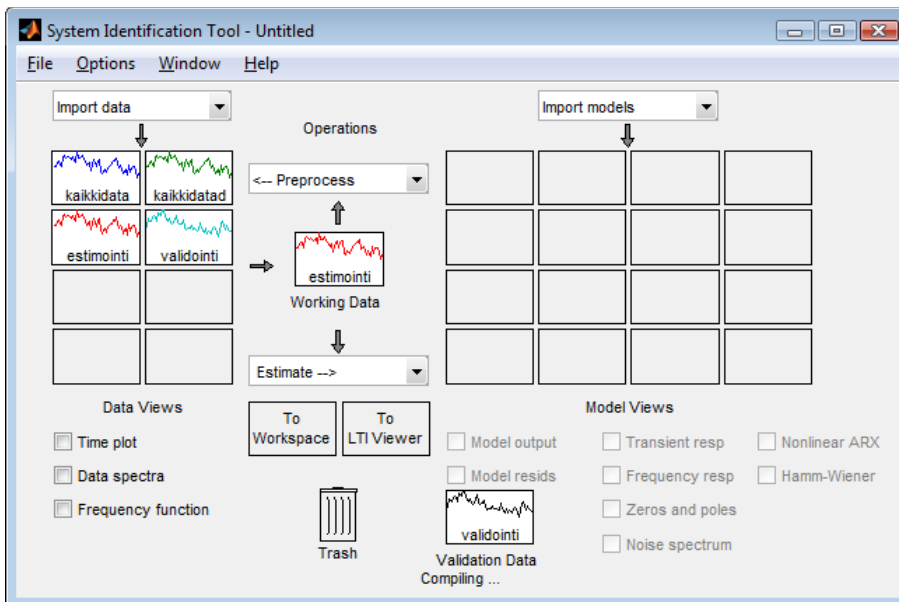
Painetaan nappulaa "Insert", jonka jälkeen pääikkunassa on koko aineiston jälkipuolisko nimellä "validointi".

Voimme varmistua siitä, että jakomme onnistui, piirtämällä aikasarjat estimointi-ja validointiaineistosta. Voit valita aktiiviset aineistot klikkaamalla niitä työkalun pääikkunassa. Kun rastitat sen jälkeen ruudun "Time plot", piirtää työkalu aikasarjat vain niistä aineistoista, jotka ovat aktiivisina. Lopputuloksen pitäisi näyttää Kuvan 6 kaltaiselta.



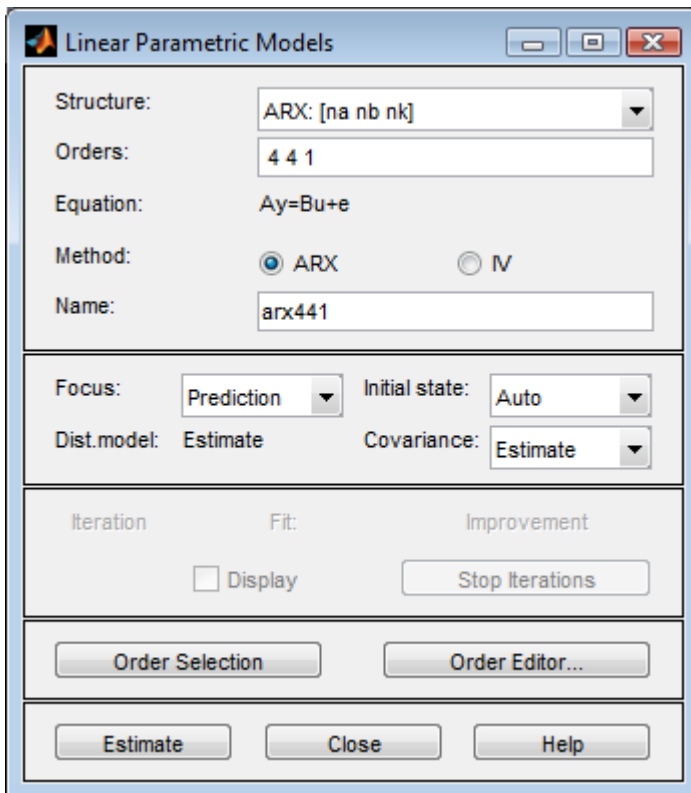
Kuva 6: Aikasarja estimointi- ja validointiaineistosta. Estimointiaineisto on esitetty punaisella ja validointiaineisto syaanilla värillä.

Koska mallit halutaan rakentaa estimointiaineiston perusteella, raahataan se "Working data"-laatikkoon. Validointiaineisto raahataan "Validation data"-laatikkoon. Nyt olemme valmiit estimoimaan malleja, ja työkalun pääikkunan pitäisi näyttää Kuvan 7 kaltaiselta.



Kuva 7: Työkalun pääikkuna, estimointi- ja validointiaineistot paikallaan.

Mallit estimoidaan valitsemalla "Working data"-laatikon alapuolelta "Estimate"-valikosta vaihtoehdon "Linear parametric models". Esiin tulee Kuvan 8 mukainen ikkuna.



Kuva 8: Työkalun ikkuna, jonka avulla voidaan estimoida

Työkalussa käytettävistä malleista saa lisää tietoa kirjoittamalla MATLAB:in komentoikkunaan

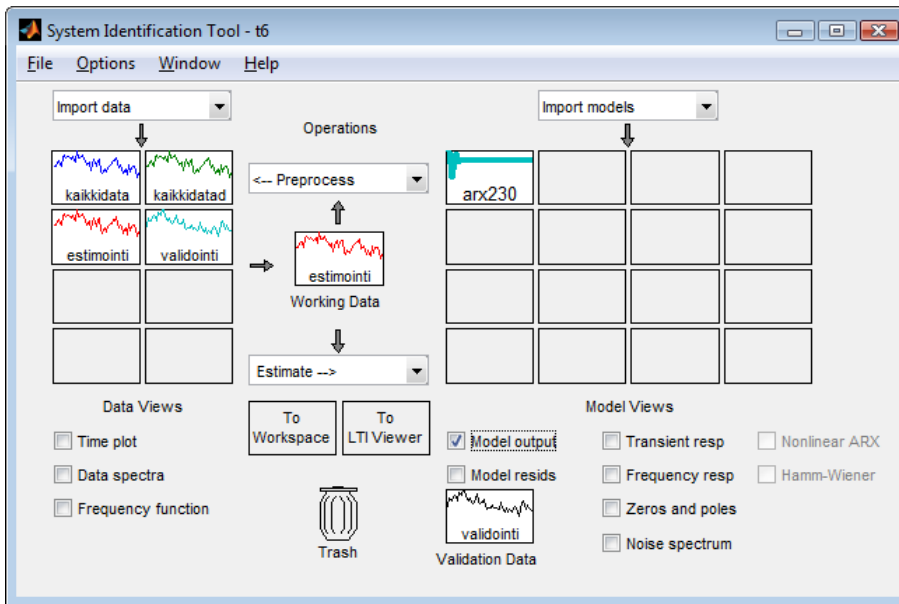
```
doc arx
doc armax
doc oe
doc bj
```

Esimerkiksi työkalun sisältämä ARX-malli on muotoa

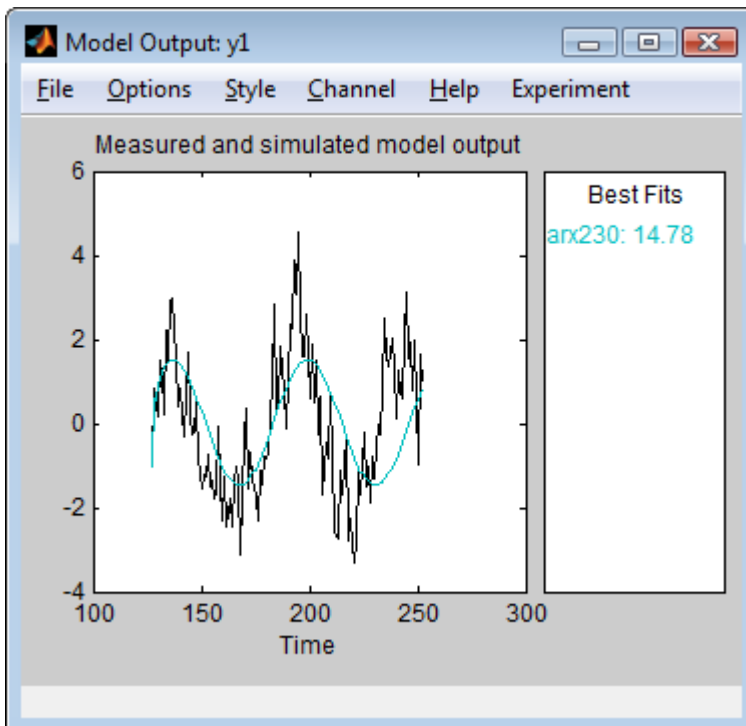
$$y(t) + a_1y(t - 1) + \dots + a_{na}y(t - na) = b_1u(t - nk) + \dots + b_{nb}u(t - nb - nk + 1) + e(t),$$

jossa  $na$ ,  $nb$  ja  $nk$  kertovat mallin kertaluvun. Nämä kertaluvut syötetään myös ARX-mallia estimoidaessa ikkunan "Orders"-laatikkoon. Valitaan esimerkin vuoksi kertaluvut  $na = 2$ ,  $nb = 3$ ,  $nk = 0$  ja painetaan "Estimate"-nappulaa. Mallin parametrit estimoidaan käyttäen "Working data"-laatikossa olevaa aineistoa, ja valmis malli ilmestyy työkalun pääikkunan oikealle puolelle kuten Kuvassa 9 on esitetty.

Rastittamalla mallien alapuolelta ruudun "Model output" piirtää työkalu aikasarjan validointiaineistosta sekä valittuna olevan mallin tekemän ennusteen. Näitä vertaamalla voidaan tehdä päätelmiä estimoinnin hyvyydestä. Esimerkki tästä on annettu Kuvassa 10.



Kuva 9: Työkalun pääikkuna, johon on lisätty ARX-malli.



Kuva 10: Aikasarja validointiaineistosta sekä ARX-mallin antama ennuste.

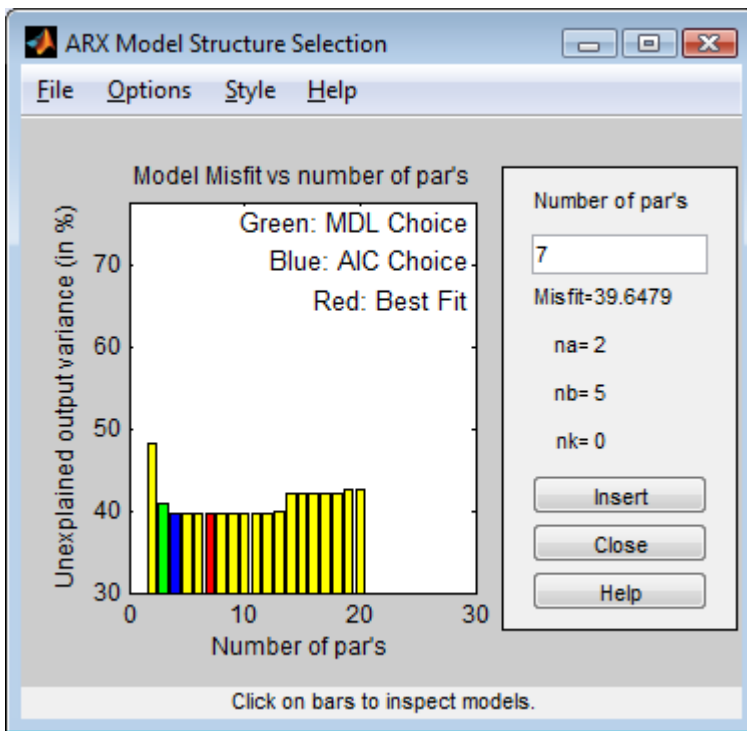
ARX-mallin tapauksessa mallin kertaluvussa voidaan käyttää hyväksi myös työkalun "Order selection"-ominaisuutta. Painamalla Kuvassa 8 esitettyssä ikkunassa "Order selection"-nappulaa, ilmestyy "Orders"-laatikkoon teksti

[0:10 0:10 0:10]

Teksti tarkoittaa sitä, että työkalu käy läpi kaikki kertalukujen kombinaatiot välillä  $na \in \{0,1, \dots, 10\}$ ,  $nb \in \{0,1, \dots, 10\}$ ,  $nk \in \{0,1, \dots, 10\}$ . Valitaan esimerkin vuoksi  $nk = 0$  eli muutetaan teksti muotoon

[0:10 0:10 0]

Painamalla "Estimate"-nappulaa saadan esille Kuvan 11 ikkuna.



Kuva 11: Mallin kertaluvun valinta.

Jokainen palkki kuvassa vastaa tiettyä mallirakennetta, ja palkin korkeus kuvastaa sitä osaa ulostulon varianssista, jota malli ei selitä. Lisäksi palkit on värikoodattu siten, että vihreä palkki kuvastaa mallia, joka tulee valituksi noudattamalla Rissanen MDL-informaatiokriteeriä, sininen noudattamalla AIC-informaatiokriteeriä ja punainen palkki kuvastaa aineistoon parhaiten sopivaa mallia. Mallinvalintakriteereistä kerrotaan lisää myöhemmässä laskuharjoituksessa sekä luentokalvoissa 10.

Valitaan esimerkiksi MDL:n ehdottama mallirakenne klikkaamalla vihreää palkkia, jonka jälkeen mallin parametrien lukumäärä sekä eri kertaluvut tulevat ikkunaan näkyviin. Painamalla "Insert"-nappulaa malli siirtyy suoraan estimointityökaluun.

ARMAX, OE ja BJ-mallit voidaan estimoida samalla tavalla kuin ARX-malli valitsemalla Kuvassa 8 esitetyssä ikkunassa "Structure"-pudotusvalikosta haluttu rakenne. Huomaa myös, että estimoidusta mallista saat lisää tietoa (esim. kertoimet) klikkaamalla siihen oikealla napilla työkalun pääikkunassa.

Kuvassa 7 esitetyssä vaiheessa oleva työskentelytila löytyy tehtävän tiedostoista nimellä "t6.sid". Voit avata sen valitsemalla pääikkunan "File"-valikosta vaihtoehdon "Open session".