

Ympäristötilastotieteen peruskurssi syksy 2014, harjoitukset ja demonstraatiot 2

Dosentti

Jukka Hoffrén

Helsingin yliopisto, Tilastokeskus

Sosiaalitieteiden laitos (Tilastotiede), Valtiotieteellinen tiedekunta

27.11.2014

Harjoitustyö 2 op.

- Esseemuotoinen raportti, jossa sovelletaan kurssilla käytyjä menetelmiä esim. omaan tutkimusaiheeseen.
- Sovella omaan tutkimusaiheeseesi kvantitatiiviseen aineistoon jotain kurssilla esitettyä tai jotain muuta ympäristötilasto-tieteen menetelmää. Esim. ekotehokkuutta, skenaariomallintamista jne.
- Esseiden laajuus 6-8 sivua (tekstiä, taulukoita ja kuvioita). 1 op.
- Palautus 3.12.2014 mennessä.
- Jos esittää viimeisellä luentokerralla työnsä saa 1 op lisää. Lähetä työ etukäteen sähköpostilla 2.12. mennessä: jukka.hoffren@helsinki.fi

Harjoitustyön sisältö

- Aiheen ja aineisto esittely
- Käytetyn menetelmän esittely
- Analyysi
- Tulosten esittely
- Tulosten analysointi

Harjoitukset 27.11.

- To 27.11. klo 15-17 SSKH IT-luokka (Snellmaninkatu 12).
- Käytettävissä SAS 9.4.

Sisältö

Harjoitukset 2:

1.3 Lineaarinen mallinnus

2.1 Epälineaarinen mallinnus

2.2 Öljyhuipun mallinnus

1.3. Lineaarinen mallinnus

$GPI = a + bX + cY + dTE + e$, missä

X on BKT (bruttokansantuote)

Y on DMI (suora materiaalien kulutus)

Z on TE (total energy)

$$\Rightarrow GPI = 5939,58785 + (-0,59923 * BKT) + (73,10062 * DMI) + (0,00103 * TE)$$

Lineaarinen regressio

```
libname harj 'f:\';  
data harj.harj3_sas;set harj.harj3_sas;  
proc gplot;plot (bkt dmi gpi)*year;run;  
proc reg;model gpi=bkt dmi te;run;
```

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	30317395	10105798	3.02	0.0440
Error	32	107039960	3344999		
Corrected Total	35	137357355			

Root MSE	1828.93377	R-Square	0.2207
Dependent Mean	8746.38889	Adj R-Sq	0.1477
Coeff Var	20.91073		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	5939.58785	3167.43532	1.88	0.0699
BKT	BKT	1	-0.59923	0.28952	-2.07	0.0466
DMI	DMI	1	73.10062	35.66858	2.05	0.0487
TE	TE	1	0.00103	0.00464	0.22	0.8255

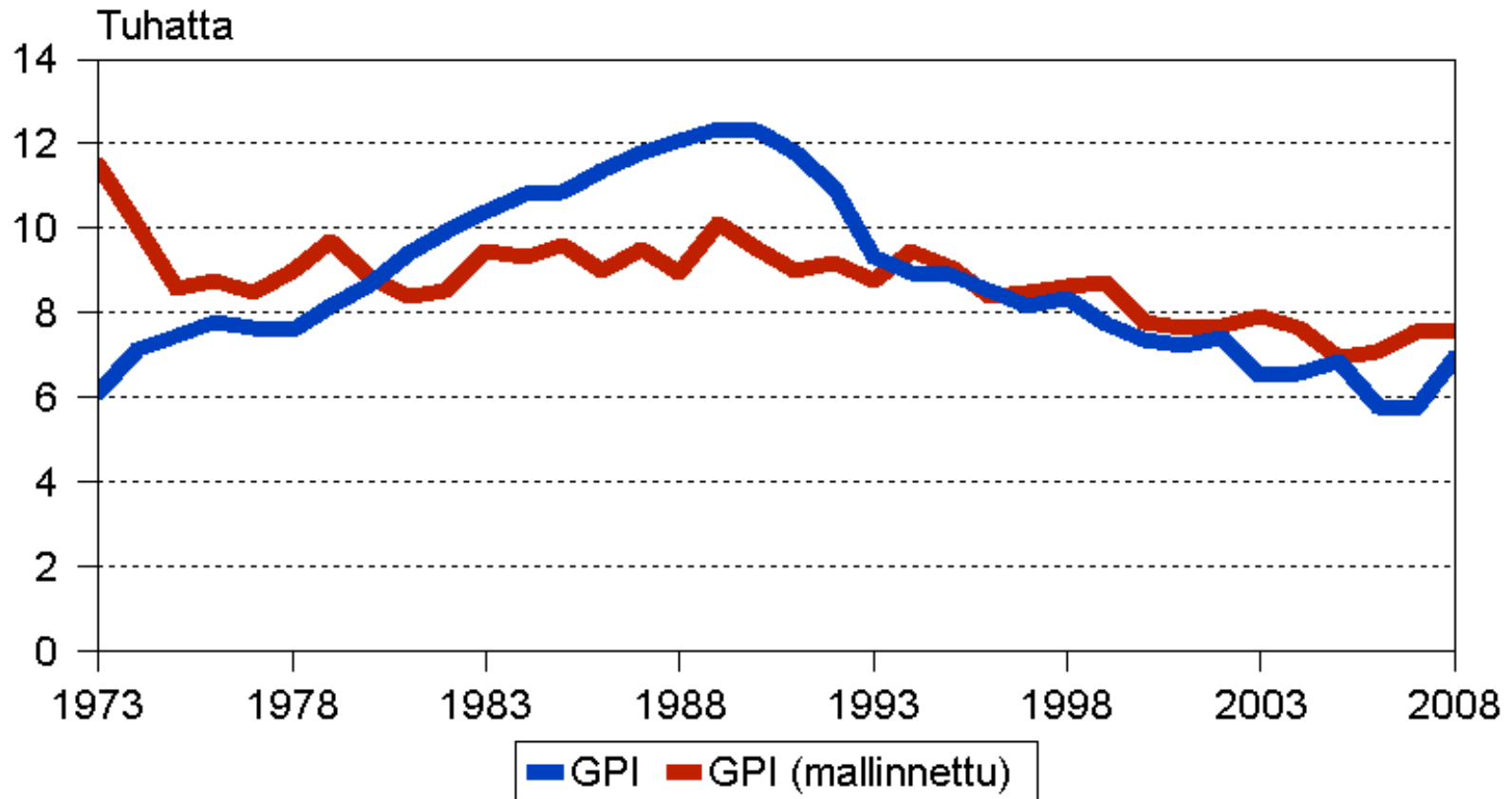
The SAS System 13:45 Thursday, March 28, 2013 4

The REG Procedure
Model: MODEL1
Dependent Variable: GPI GPI

Number of Observations Read	36
Number of Observations Used	36

Harjoitus 1.3 graafisesti

$R^2 = 22,07\%$
Adjusted $R^2 = 14,77\%$



2.1. Epälineaarinen mallinnus

$GPI = a + bX + cY + dTE + e$, missä

X on BKT (bruttokansantuote)

Y on DMI (suora materiaalien kulutus)

Z on TE (total energy)

$$\Rightarrow GPI = -15704 + (3,69674 * BKT) + (-230,78716 * DMI) + (0,01735 * TE) + (-0,00007931 * BKT^2) + (0,70904 * DMI^2) + (-0,000000013767 * TE^2)$$

Epälineaarinen regressio

```
libname harj 'f:\';  
data harj.harj3_sas;set harj.harj3_sas;  
bkt2=bkt*bkt;dmi2=dmi*dmi;te2=te*te;  
proc gplot;plot (bkt dmi gpi)*year;run;  
proc reg;model gpi=bkt dmi te bkt2  
dmi2 te2;run;
```

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	119064296	19844049	31.46	<.0001
Error	29	18293058	630795		
Corrected Total	35	137357355			

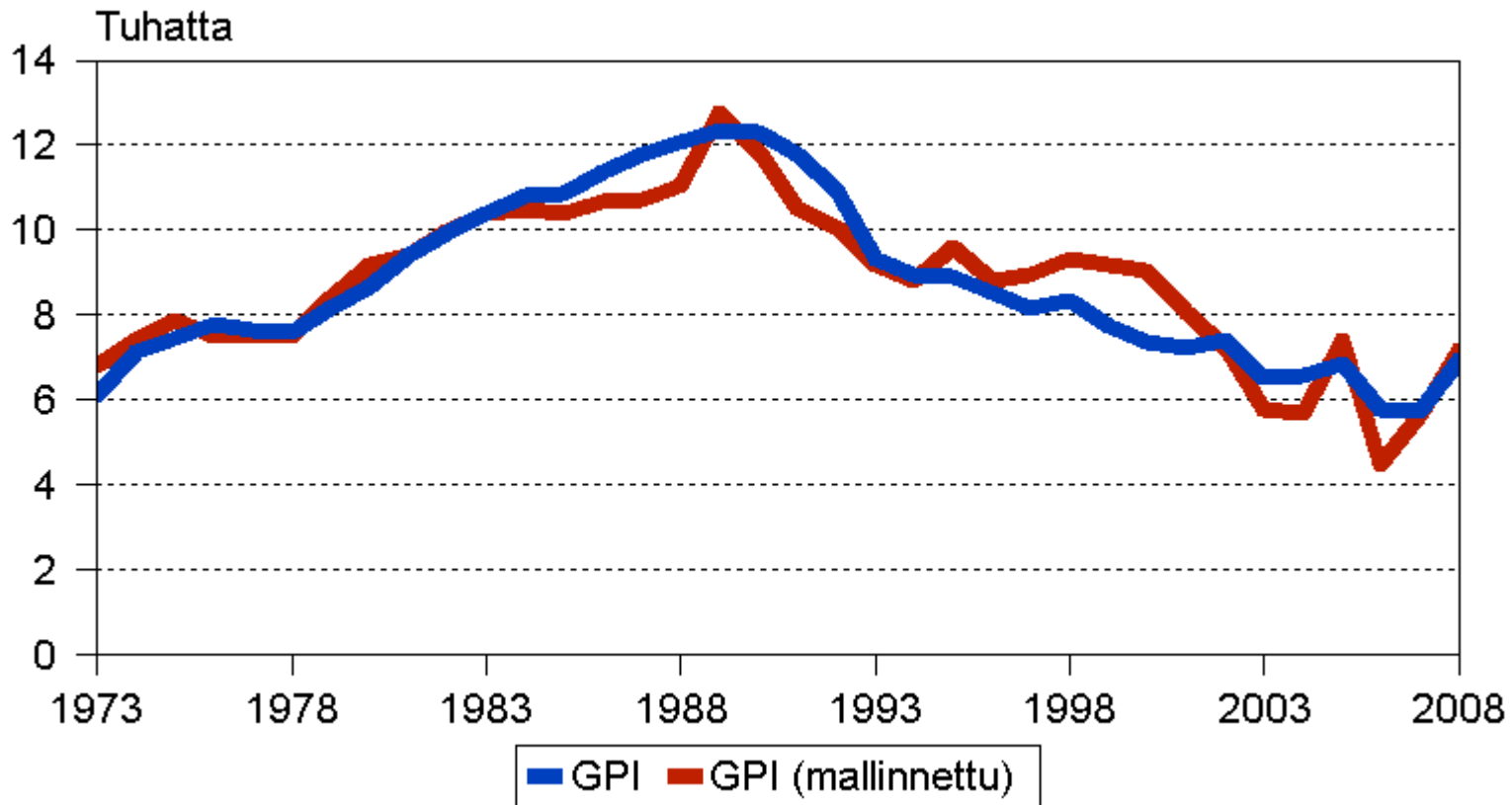
Root MSE	794.22611	R-Square	0.8668
Dependent Mean	8746.38889	Adj R-Sq	0.8393
Coeff Var	9.08062		

Parameter Estimates

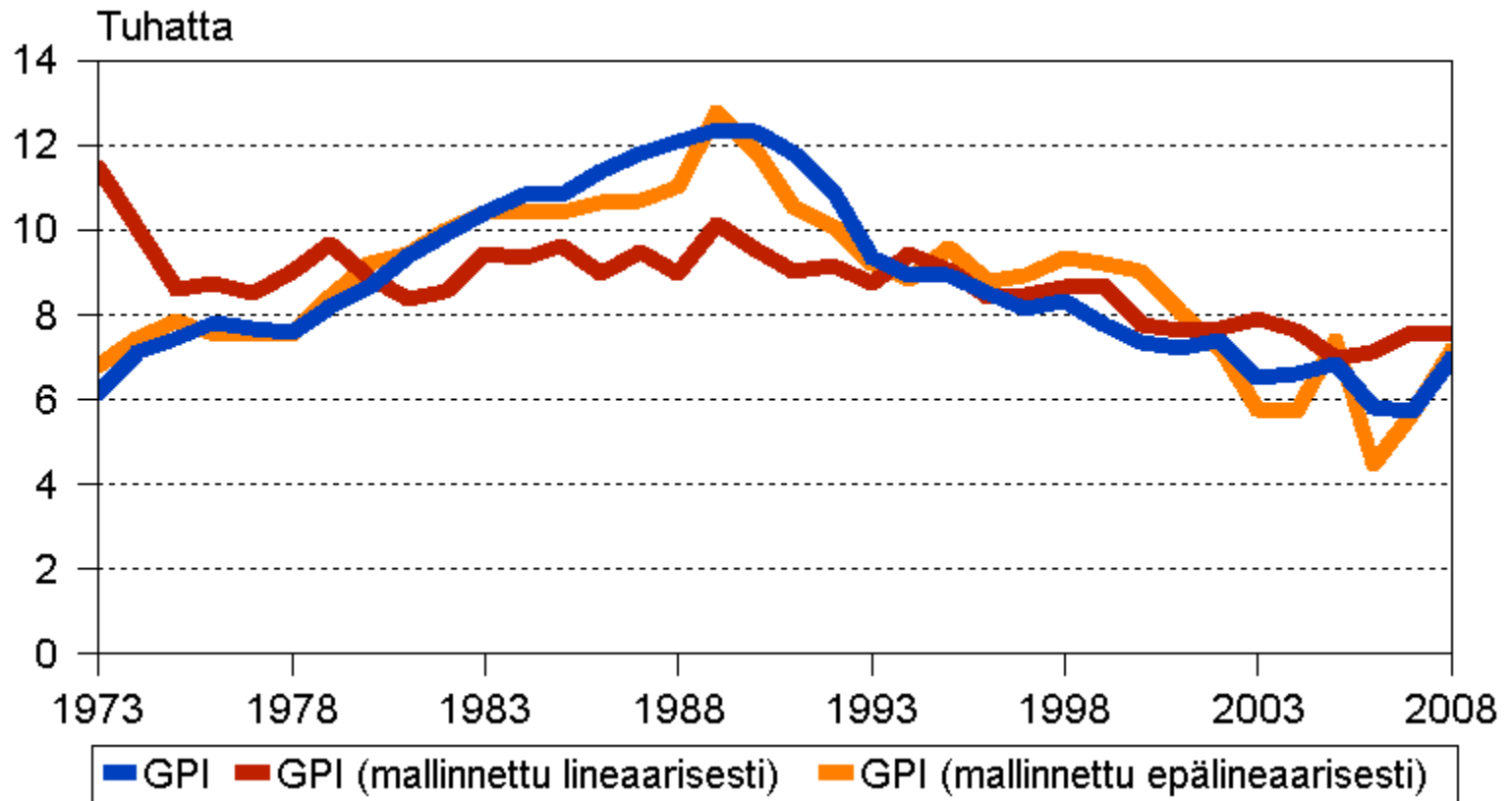
Variable	Label	Parameter		t Value	Pr > t	
		DF	Estimate			
Intercept	Intercept	1	-15704	8265.69777	-1.90	0.0674
BKT	BKT	1	3.69674	0.69804	5.30	<.0001
DMI	DMI	1	-230.78716	115.04367	-2.01	0.0543
TE	TE	1	0.01735	0.01390	1.25	0.2219
bkt2		1	-0.00007931	0.00001587	-5.00	<.0001
dmi2		1	0.70904	0.31750	2.23	0.0334
te2		1	-1.3767E-8	5.674971E-9	-2.43	0.0217

Harjoitus 2.1 graafisesti

$R^2 = 86,68\%$
Adjusted $R^2 = 83,93\%$



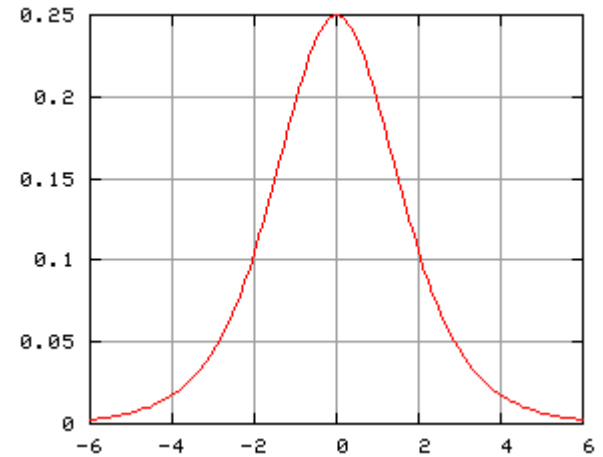
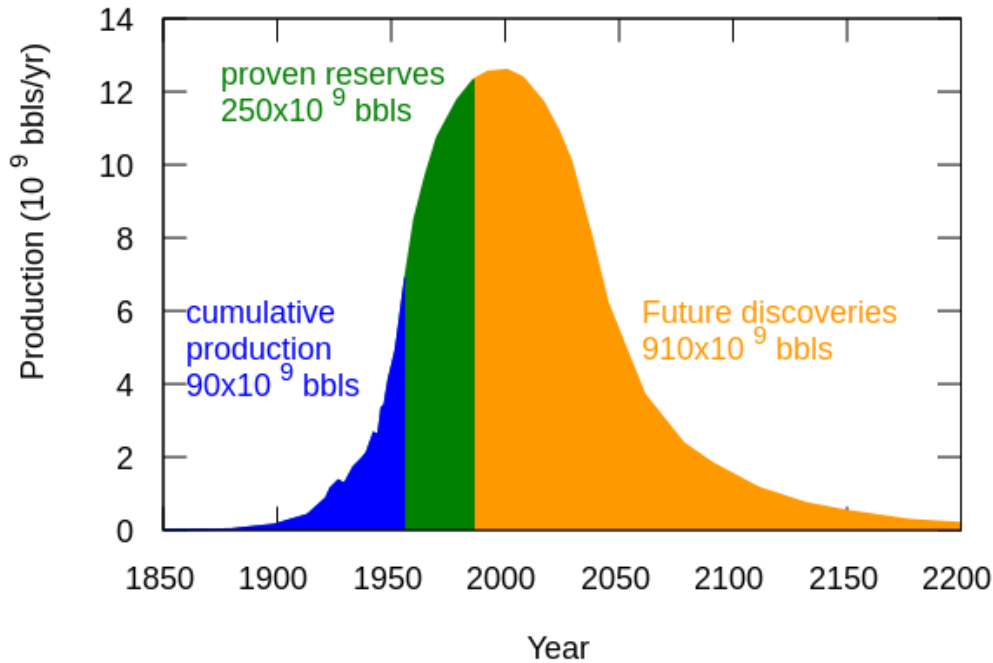
Vertailu



2.2. Öljyhuipun mallinnus

- Hubbetin käyrä on matemaattinen malli, joka on tarkoitettu mallintamaan öljyn tuotannon ja varannon määriä.
- Kehittäjä: Marion King Hubbert 1956
- Käyrä kuvaa luonnonvaran varannon hyödyntämistä ajassa (virtaa) olettaen että varantoa hyödynnetään optimaalisesti (parhainta mahdollista tekniikkaa käyttäen)
- Käyrä on luonteeltaan logaritminen
- Käyrän tarkka muoto riippuu varannon koosta ja hyödyntämisen intensiivisyydestä

Tuotantokäyrän logistinen jakautuma



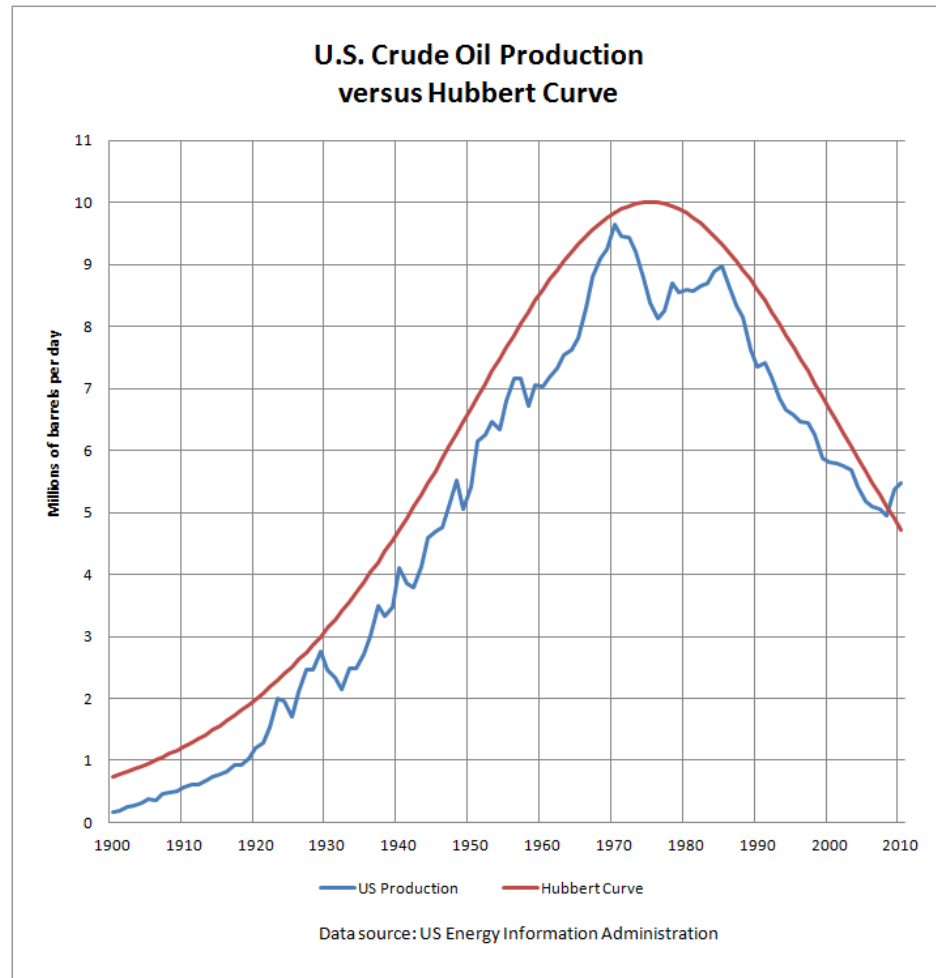
$$x = \frac{e^{-t}}{(1 + e^{-t})^2} = \frac{1}{2 + 2 \cosh t}$$

Q $Q(t)$ a b

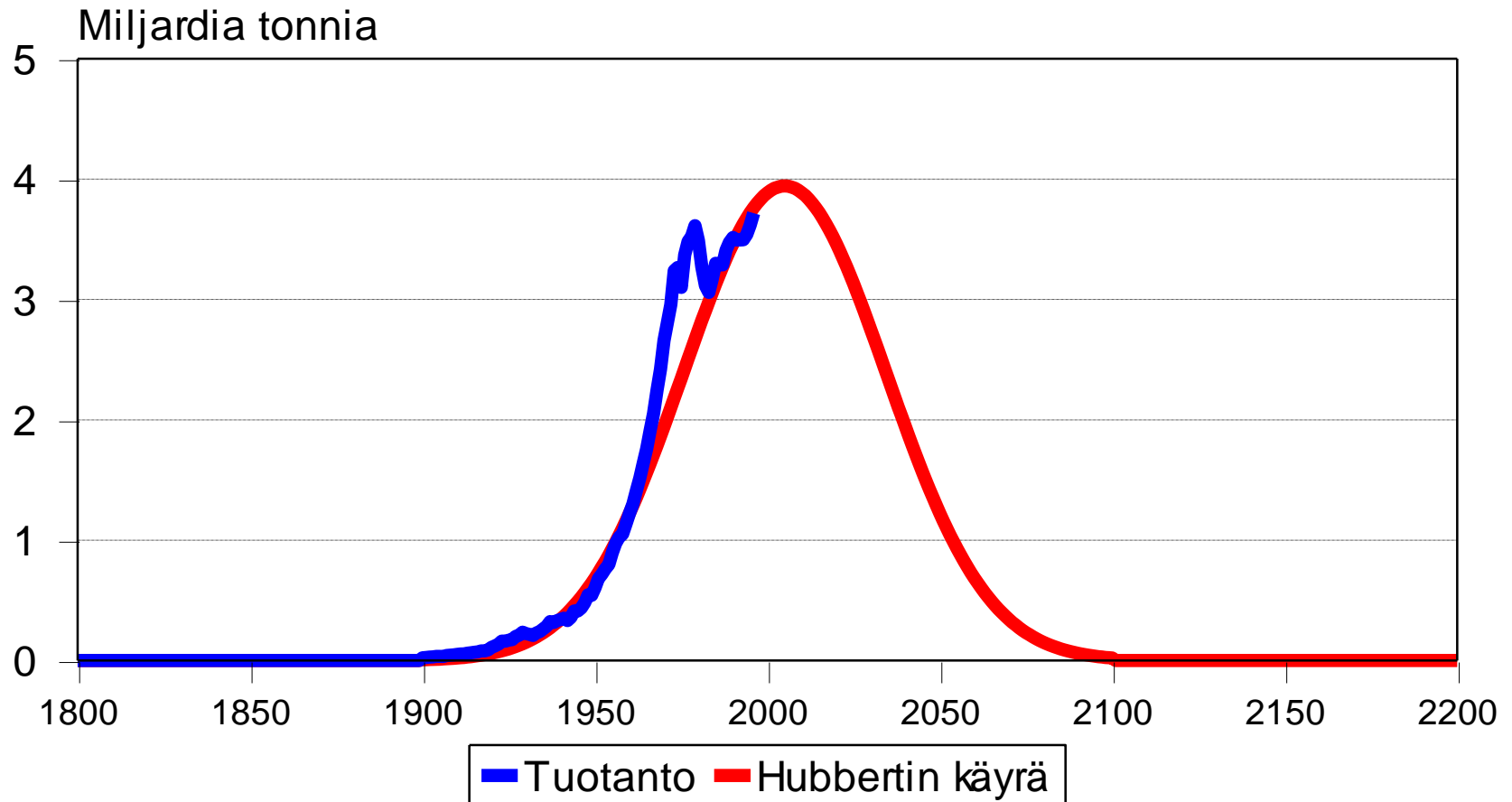
Manner-USA:n öljyntuotannon kehitys

$$Q(t) = \frac{Q_{\max}}{1 + ae^{-bt}}$$

missä Q_{\max} on käytettävissä olevan öljyn määrä, Q_t on kumulatiivinen tuotanto, a ja b ovat vakioita.



Maailman öljyntuotanto ja Hubbertin käyrä



SAS 9.4 ajovirta 1

```
libname oil 'g:\oil_production';  
/* Excelin tiedot jo luettu SAS-tiedostoon  
work.world  
ja talletetaan hakemistoon oil */
```

```
data world;set oil;  
/* muutetaan vuosilukuja pienemmiksi*/
```

```
time=year-1890;  
Qprod + actual;  
if year gt 2012 then qprod=.;  
run;
```

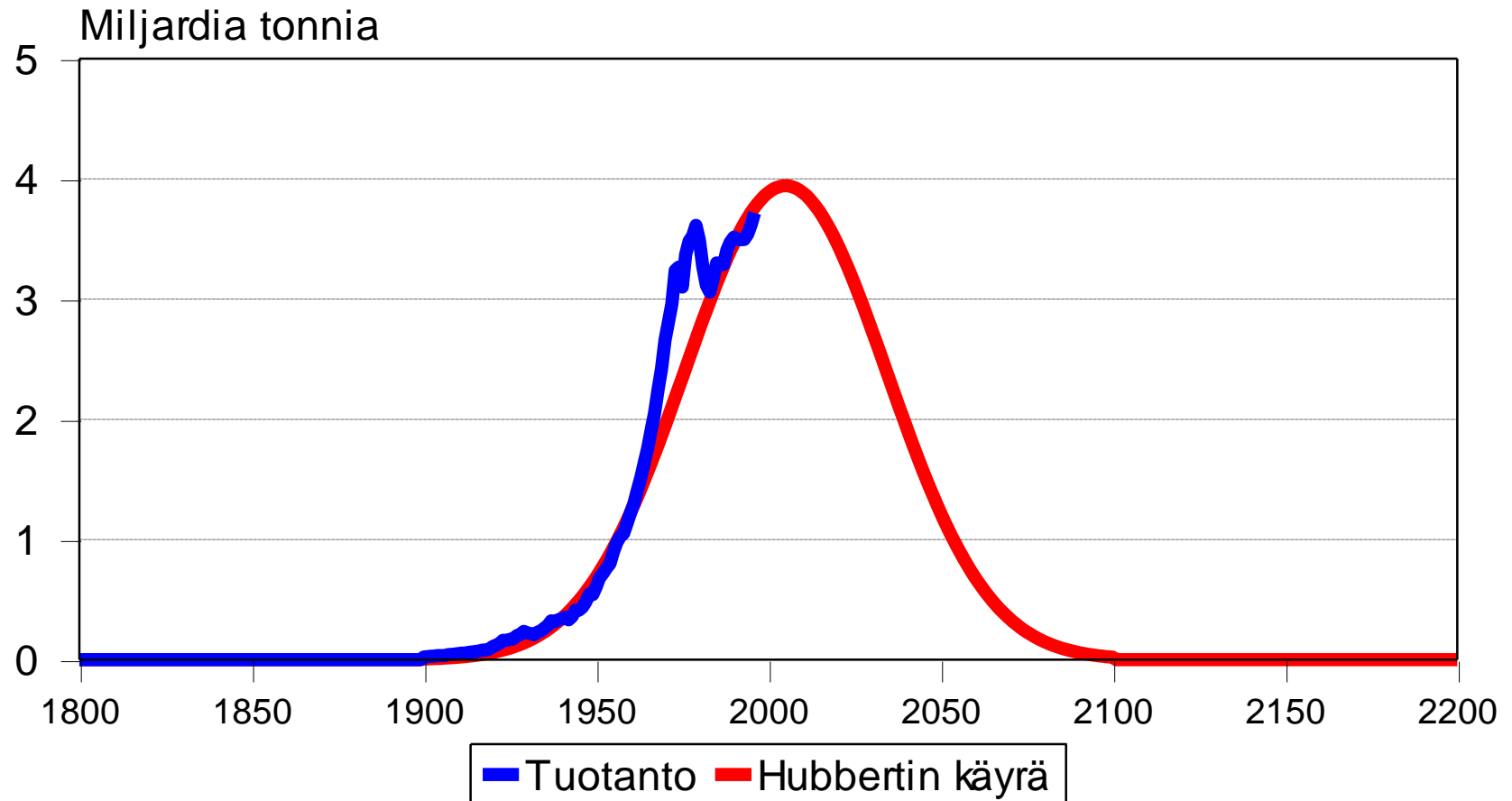
SAS 9.4 ajovirta 2

```
proc nlin data=world;  
where 1850 le year;  
/* sovitetaan tuotantoon  
Hubbertin kasvukäyrä SAS NLIN -  
proseduurilla. Parametrit saa  
normaalista SAS -tulostuksesta. Q_Max  
viimeiseltä vuodelta 2012 */  
Q_max=2*1202.872;  
Retain Q_max;  
parameters a=100 b=1;  
model actual=(Q_max*a*b*exp(-b*time))/(1  
+ a*exp(-b*time))**2;  
output out=list1 p=Hubbert;
```

SAS 9.4 ajovirta 3

```
proc print;var year actual qprod  
old_pred Hubbert;run;  
symbol1 i=join;symbol2 i=join; symbol3  
i=join;Title 'Maailman Hubbert-  
tuotantokäyrä ja todellinen tuotanto v  
2012';  
proc gplot;where year lt 2200;plot  
(actual Hubbert old_pred)*  
year/overlay;run;
```

Maailman öljyntuotanto ja Hubbertin käyrä



Kiitos!

Seuraava luento 3.12.2014. klo 15

Jukka.hoffren@helsinki.fi