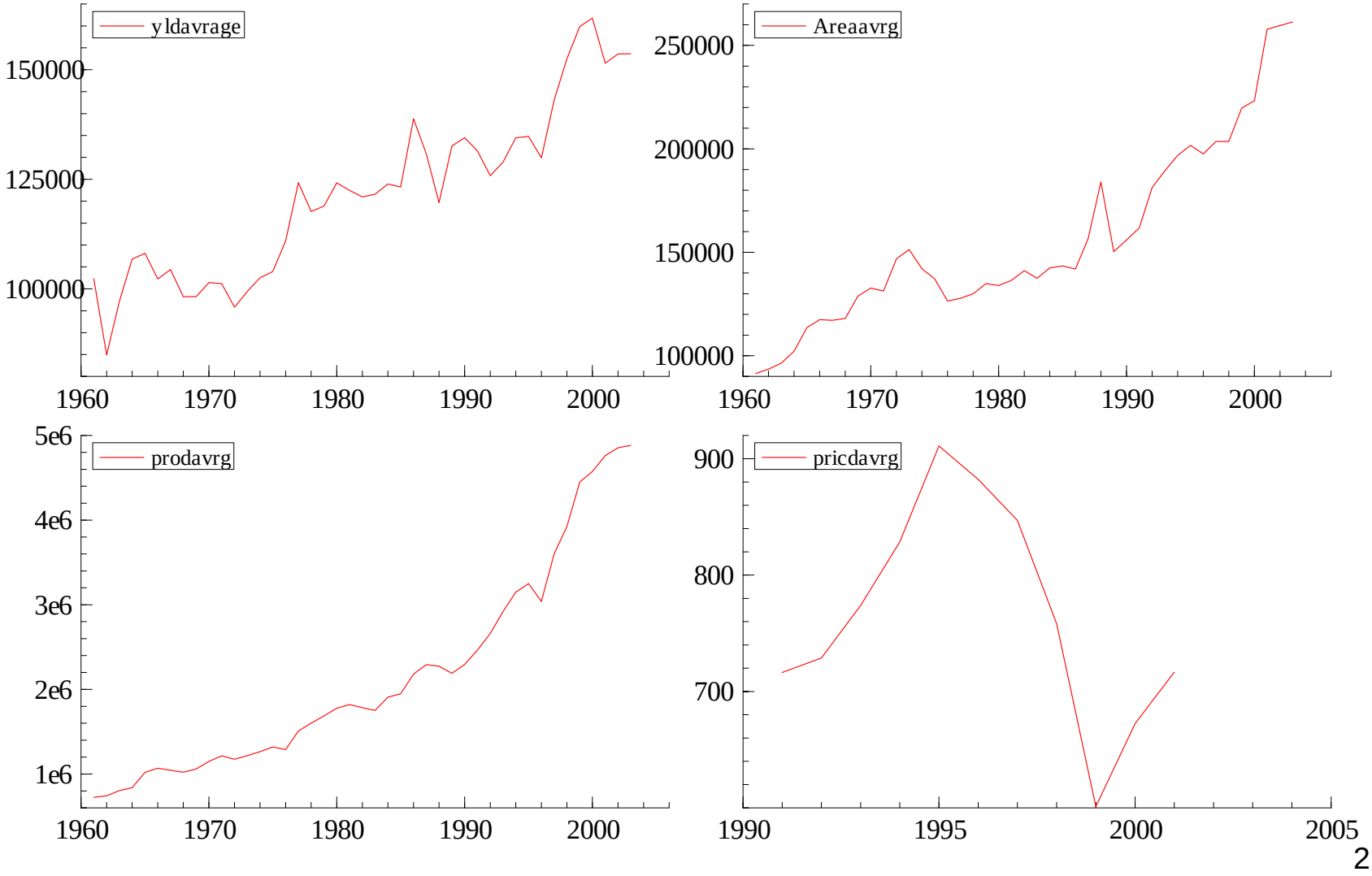
Preparation of Dataset for Quantitative Analysis

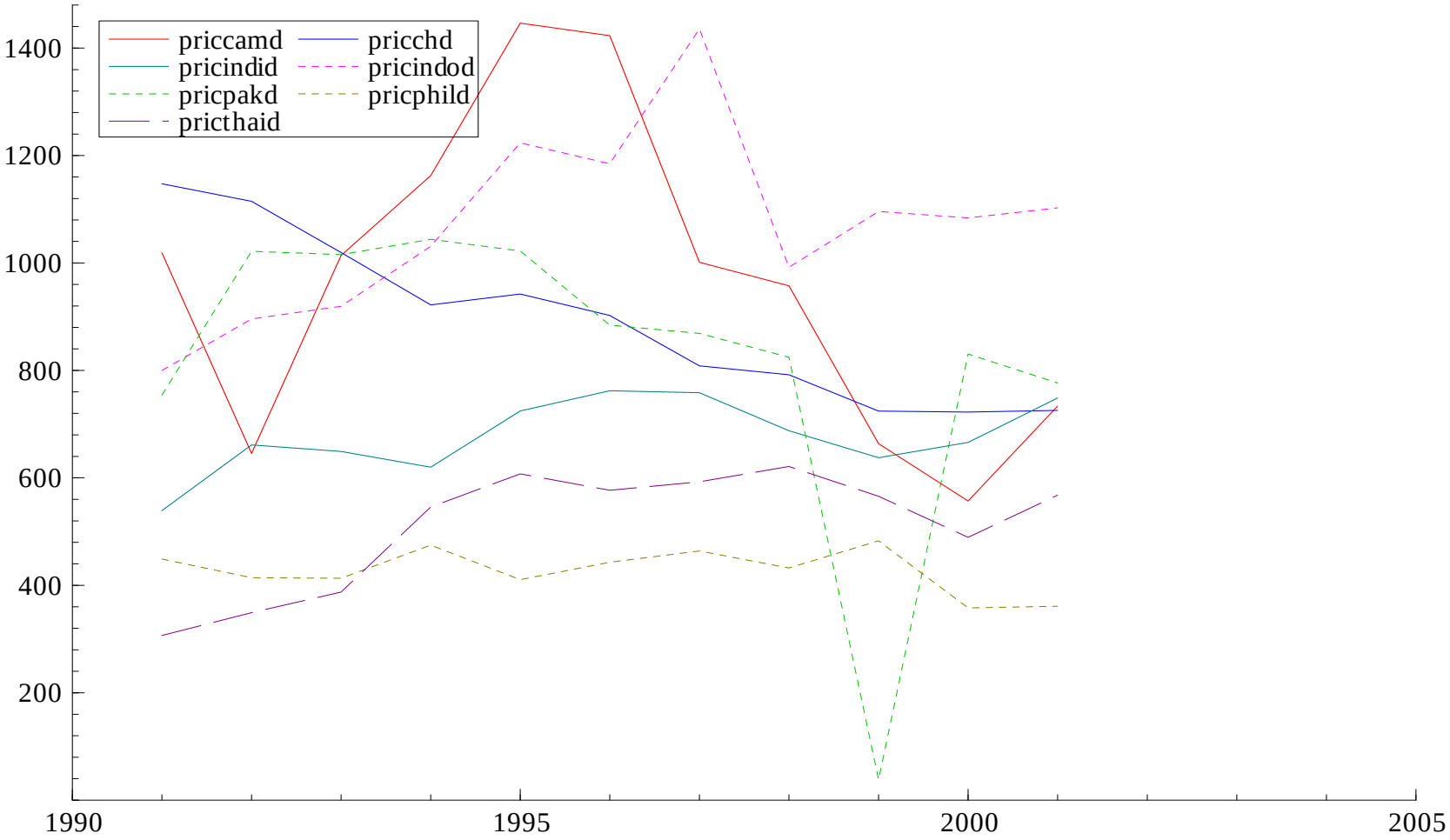
Time Series on Banana Market in Seven Asian Economies

Data Source:Food and Agricultural Organisation:
http://faostat.fao.org/faostat/collections?version=ext&hasbulk=0

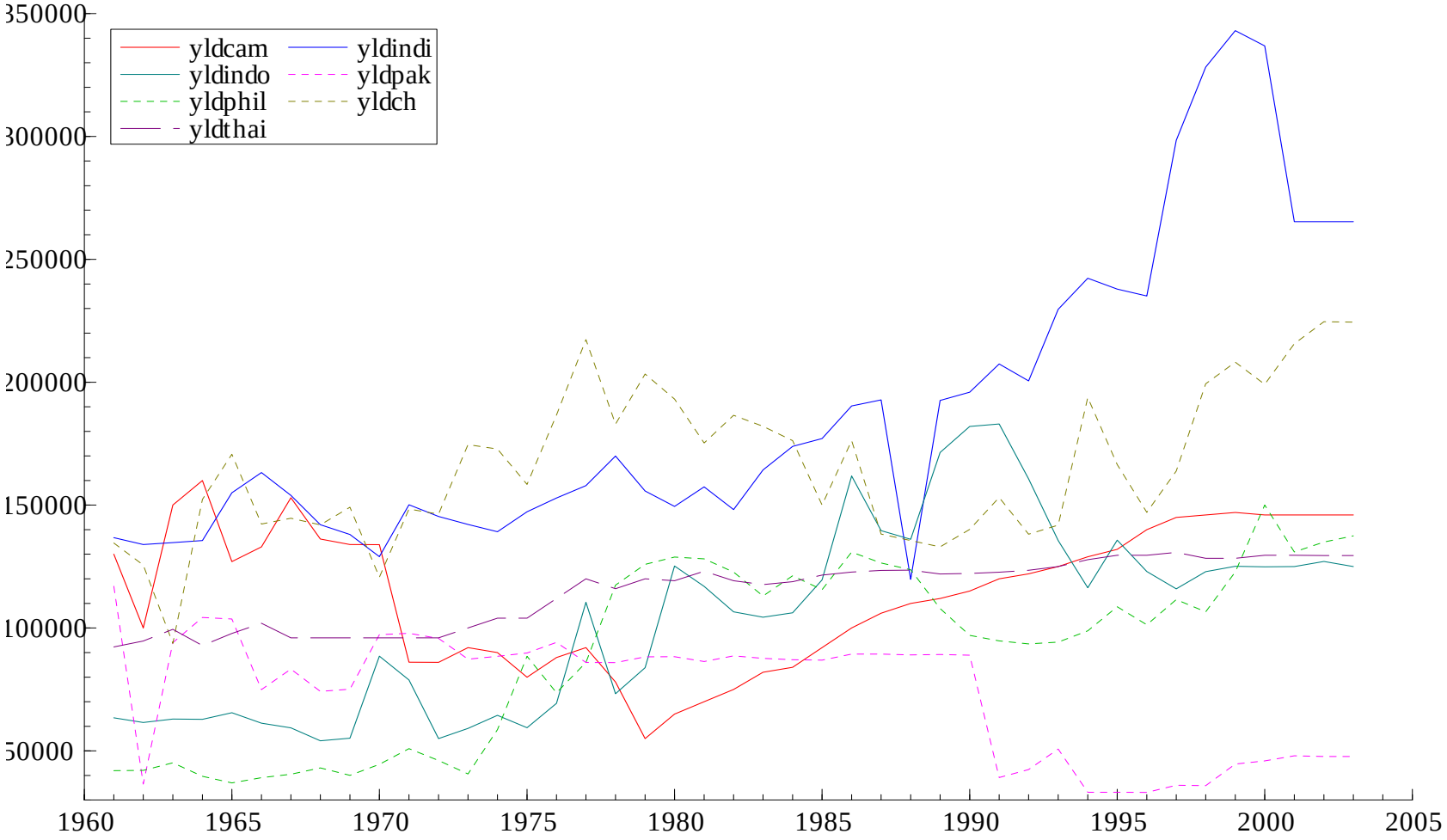
Average Yield, Area, Production and Price of Banana in Asian Countries



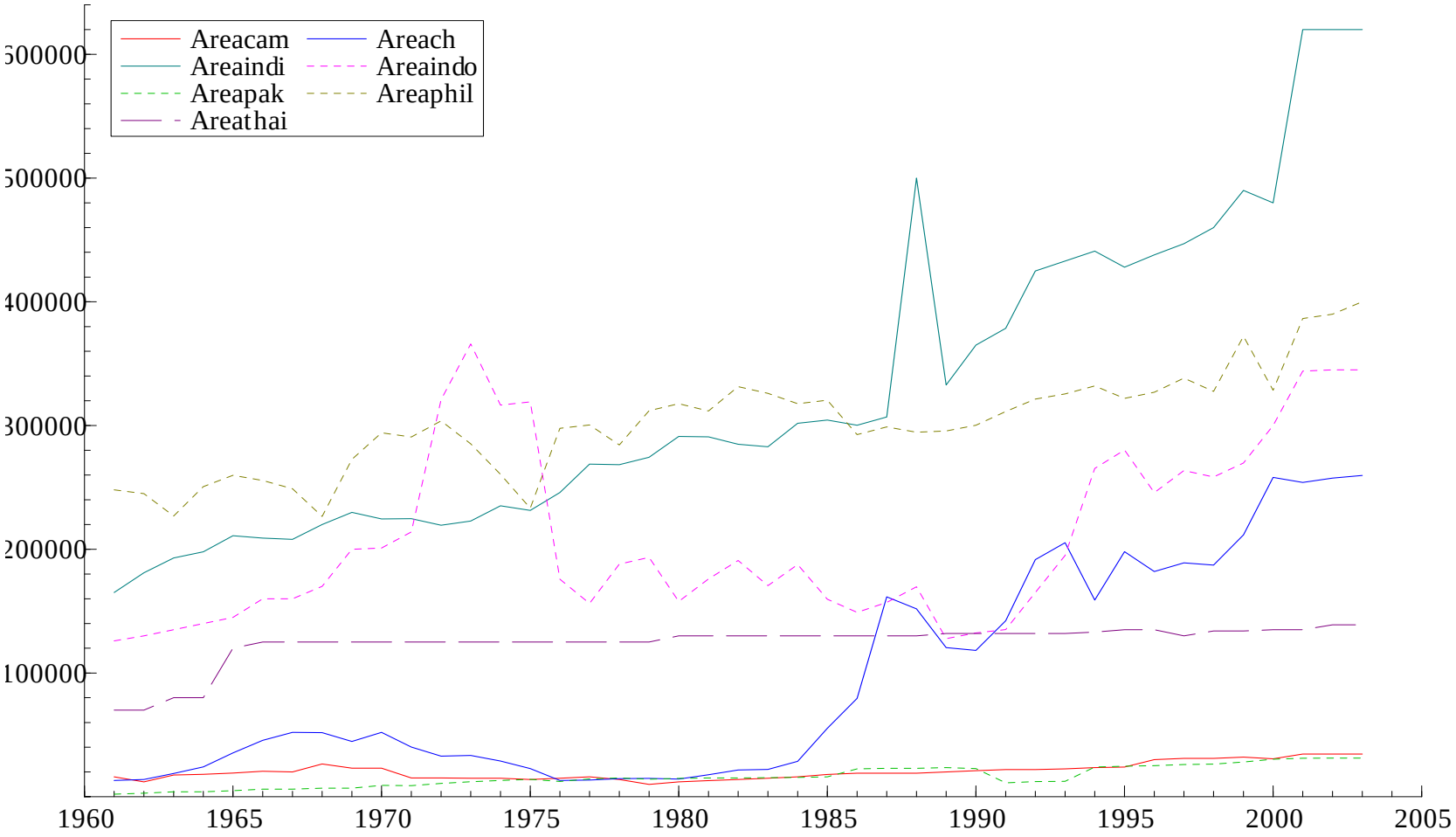
Banana Price in Asian Countries (in IMF's PPP US dollars /MT)



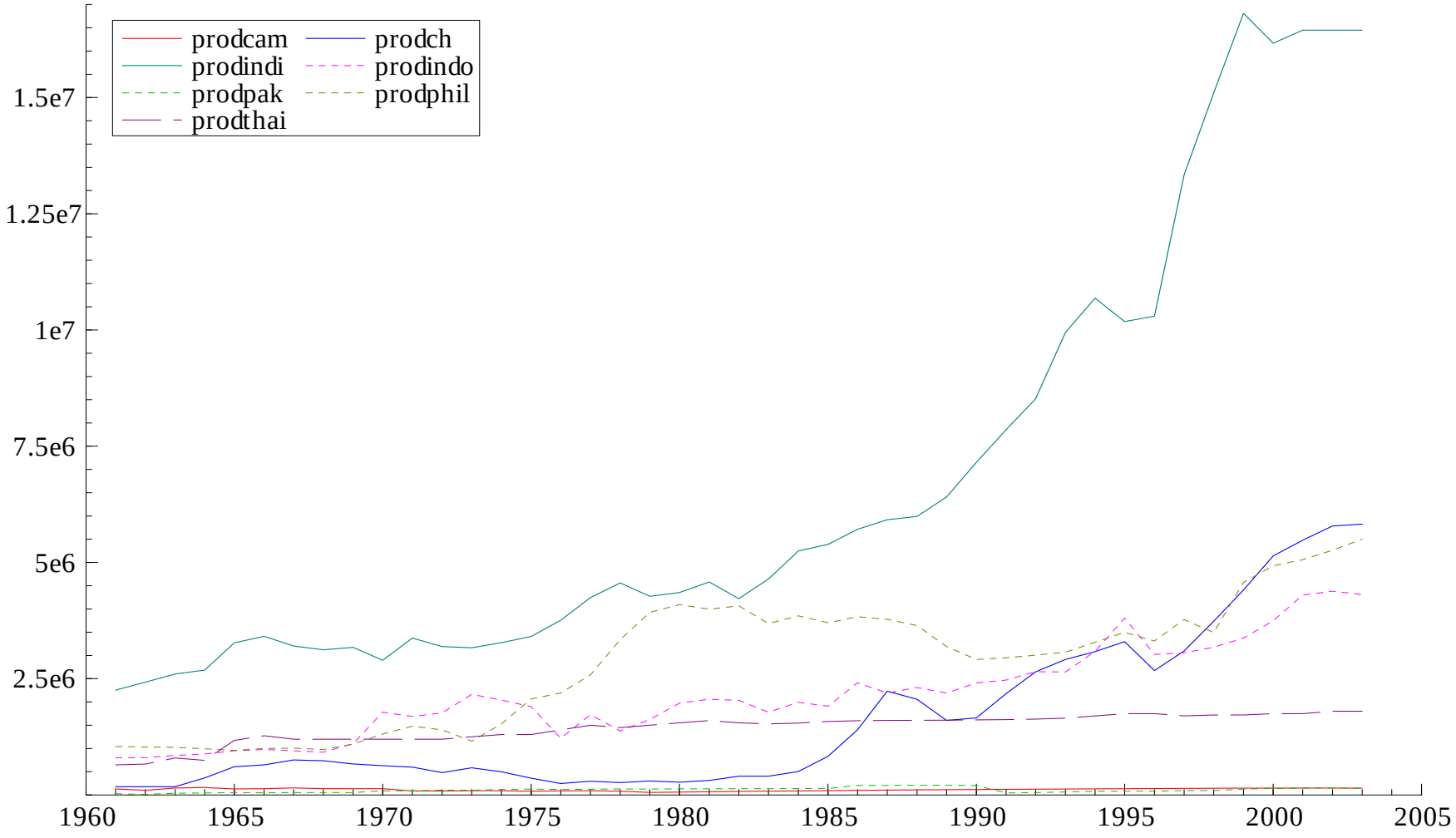
Yield of Banana Production in Asian Countries (Hg/HA)



Area of Banana Production in Asian Countries in Hectors



Production of Banana in Asian Countries (in MT)



Why can it be wrong to use raw data for regression? How can it be made right?

- Usually time series data are non-stationary.
- Using non-stationary data results in spurious regression- may result in false statement.
- Here production, area and yield of banana are all non-stationary (clear from the unit root test).
- They are non-stationary even with
  - In the first differences
  - In the log
- Only the first difference of logs was stationary

Unit Root Test of Banana in the first difference in logs

prodcam_L1: ADF tests (T=39, Constant; 5%=-2.94 1%=-3.61)

D-lag	t-adf	beta Y_1	sigma	t-DY_lag	t-prob	AIC	F-prob
2	-3.181*	0.099137	0.04863	-1.807	0.0794	-5.950	
1	-6.235**	-0.26798	0.05013	2.123	0.0407	-5.912	0.0794
0	-6.060**	0.0077992	0.05246			-5.846	0.0266

prodch_L1: ADF tests (T=39, Constant; 5%=-2.94 1%=-3.61)

D-lag	t-adf	beta Y_1	sigma	t-DY_lag	t-prob	AIC	F-prob
2	-3.564*	0.33520	0.08494	0.2420	0.8102	-4.835	
1	-4.104**	0.35912	0.08382	0.7727	0.4447	-4.884	0.8102
0	-4.426**	0.42529	0.08336			-4.919	0.7282

Prodcam_L1: production in Cambodia in the first difference
Prodch_L1: production in China in the first difference

Example of a batch file in PcGive for OSL of log difference of banana Production on the first difference on long of Area of the banana production

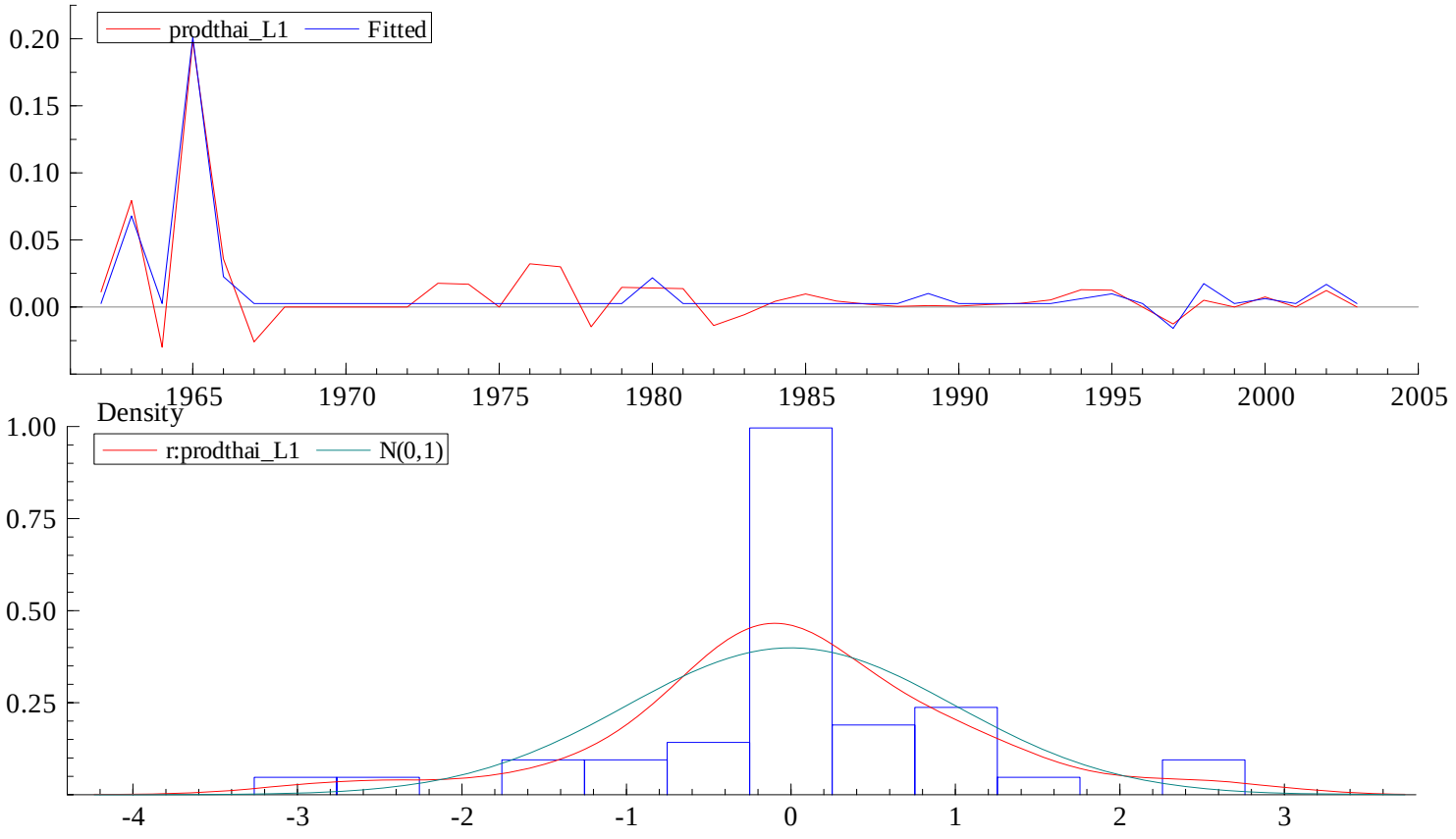
```
module("PcGive");
package("PcGive");
usedata("banana_pryaroasia_gw.xls");
system
{
    Y = prodcam_L1;
    Z = Constant, Areacam_L1;
}
estimate("OLS-CS", 1962, 1, 2003, 1);
```

```
module("PcGive");
package("PcGive");
usedata("banana_pryaroasia_gw.xls");
system
{
    Y = prodch_L1;
    Z = Constant, Areach_L1;
}
estimate("OLS-CS", 1962, 1, 2003, 1);
```

Regression of growth rate of Banana Output on Area of Production in Cambodia and Estimated Errors

prodthai_L1 = + 0.002603 + 1.127*Areathai_L1

(SE) (0.00191) (0.0654)



Errors, R-Square and Test Statistics

sigma	0.0119916	RSS	0.00575196643
R^2	0.881111	F(1,40) =	296.4 [0.000]**
log-likelihood	127.218	DW	2.14
no. of observations	42	no. of parameters	2

When the log-likelihood constant is NOT included:

AIC	-8.80065	SC	-8.71790
HQ	-8.77032	FPE	0.000150647

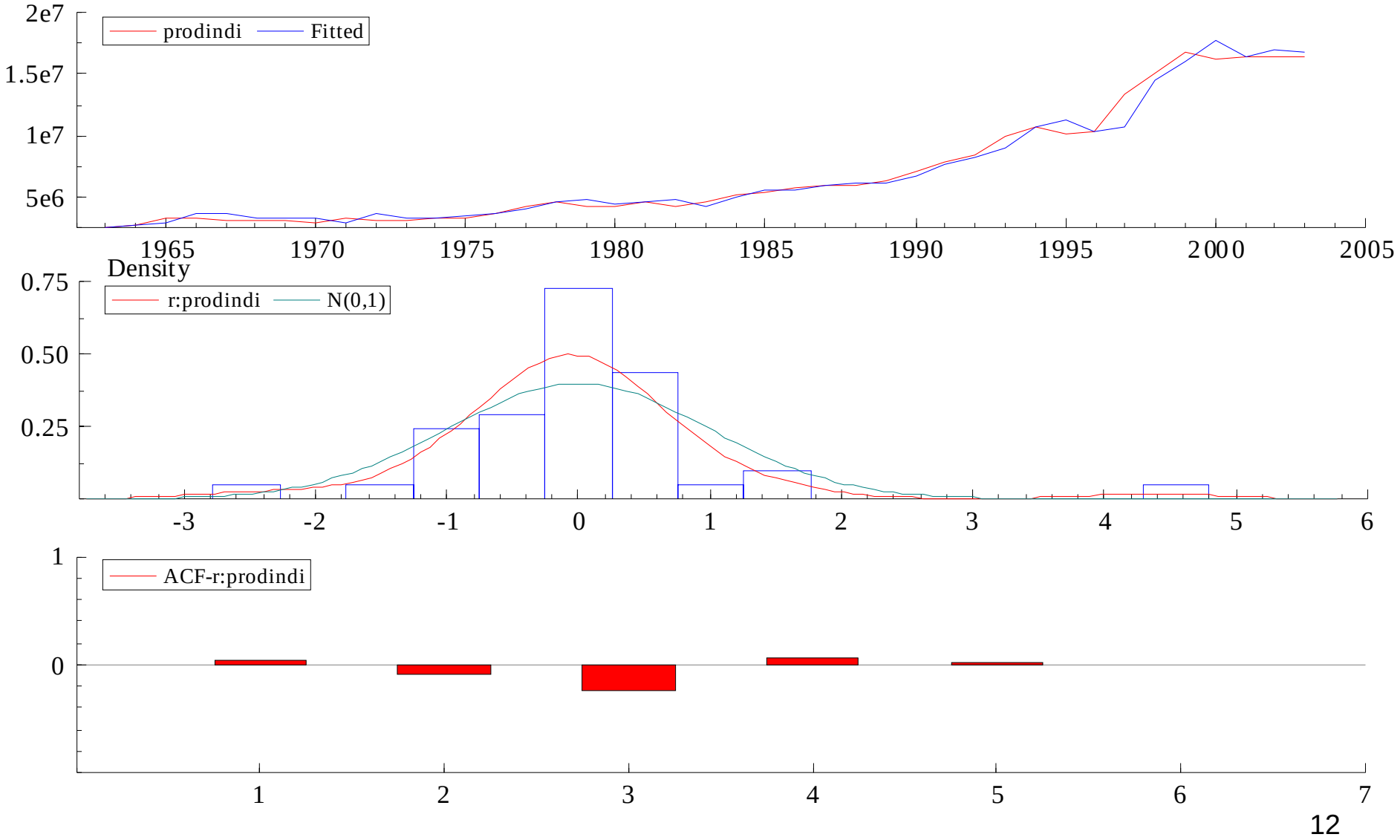
When the log-likelihood constant is included:

AIC	-5.96277	SC	-5.88002
HQ	-5.93244	FPE	0.00257297

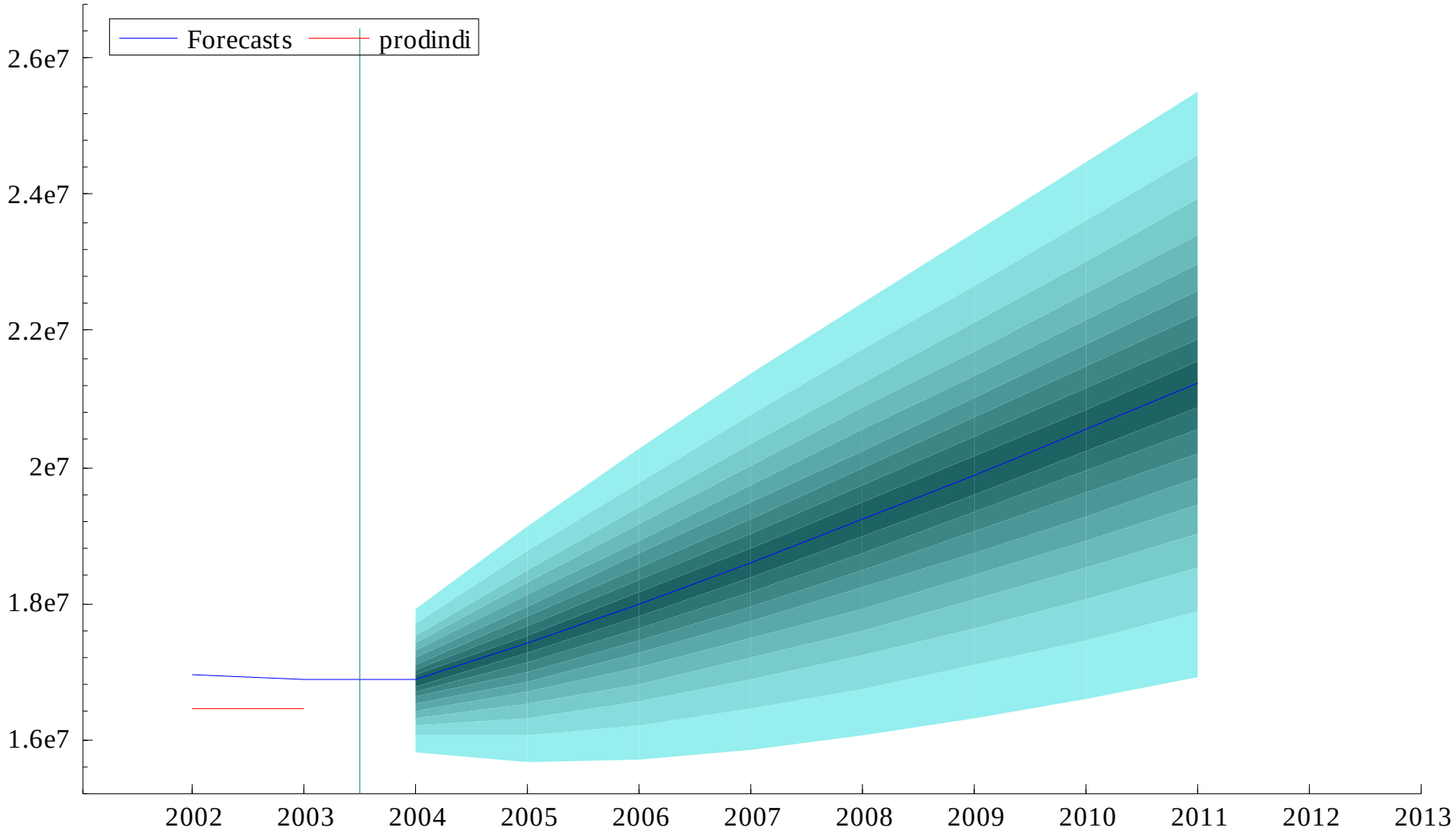
mean(prodthai_L1) 0.0105962 var(prodthai_L1) 0.00115193

Normality test: Chi^2(2) = 8.4029 [0.0150]*
hetero test: F(2,37) = 0.12773 [0.8805]
hetero-X test: F(2,37) = 0.12773 [0.8805]
RESET test: F(1,39) = 0.20695 [0.6517]

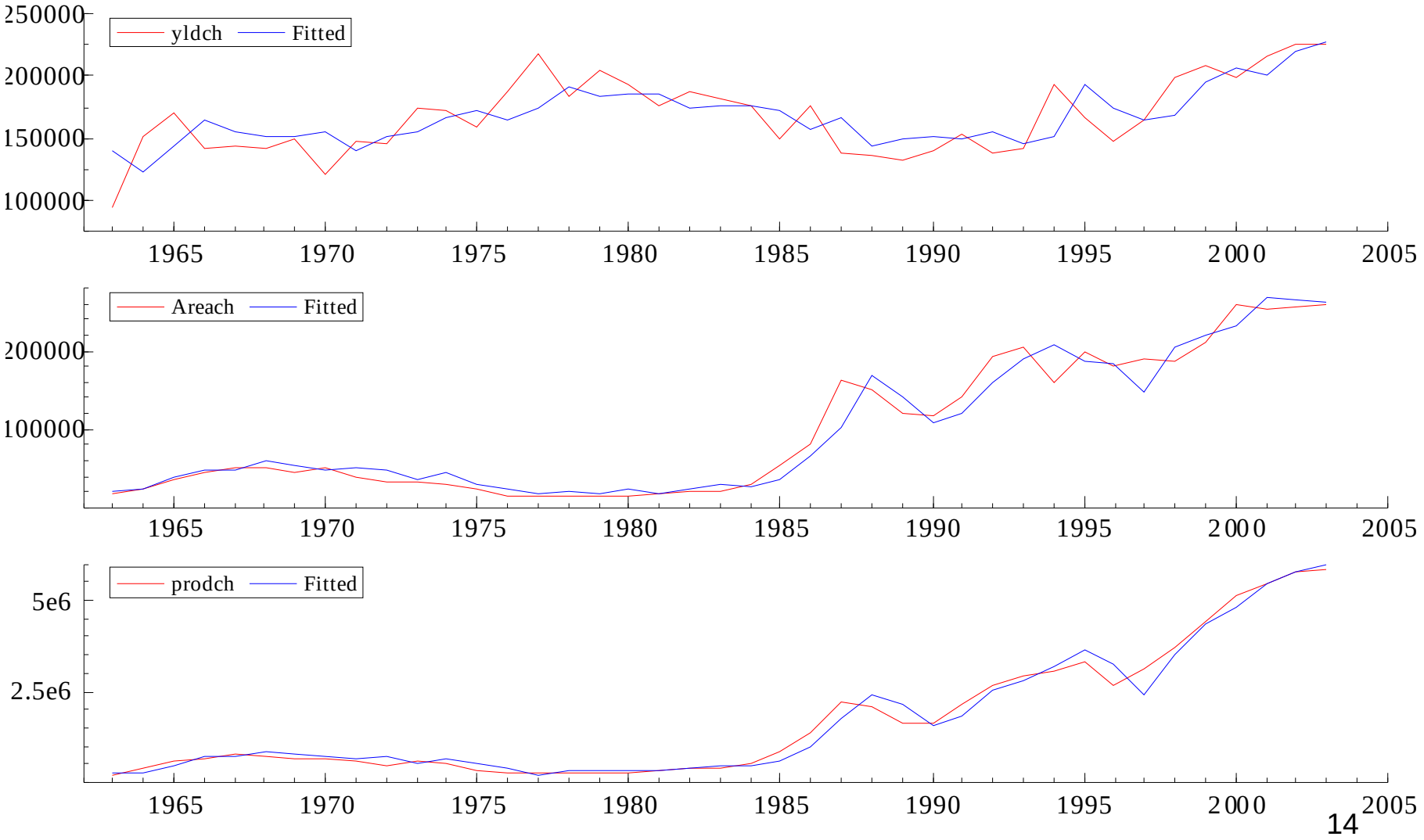
AR(2) Model of Banana Production in India



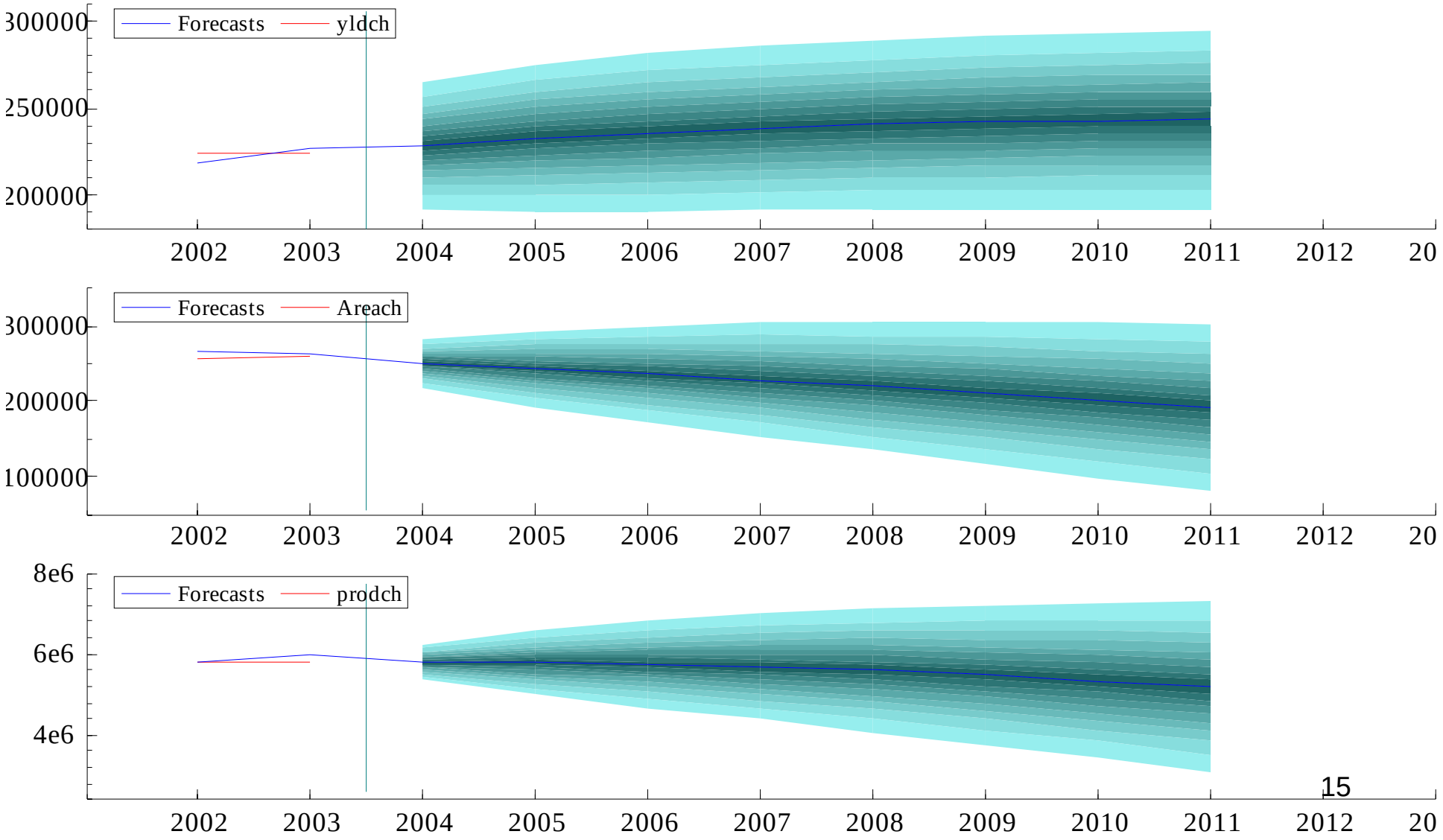
Forecast of Banana Production in India using AR(2) Model



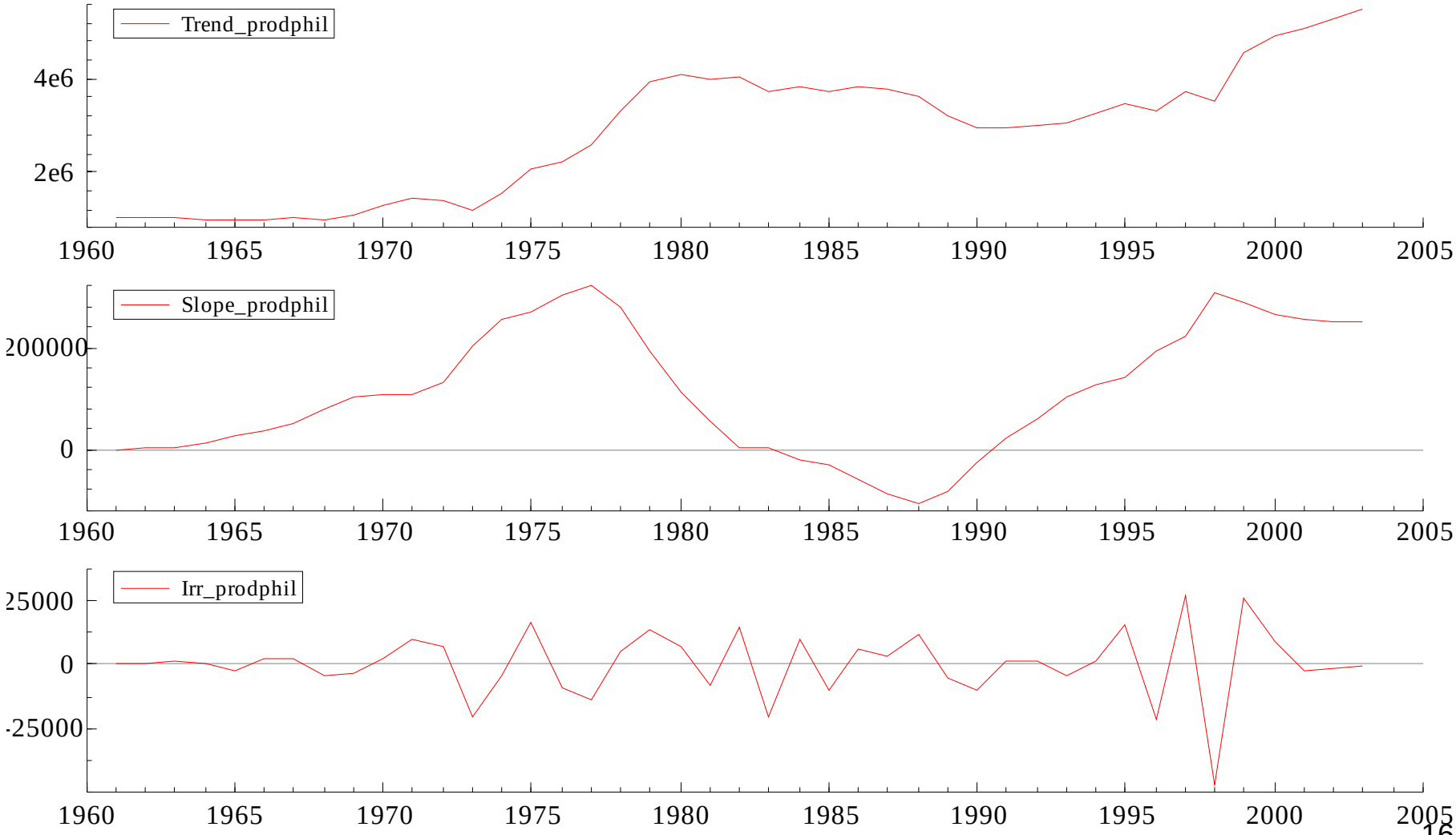
Historical simulation with Simultaneous Equation Model of Area, Production and Yield



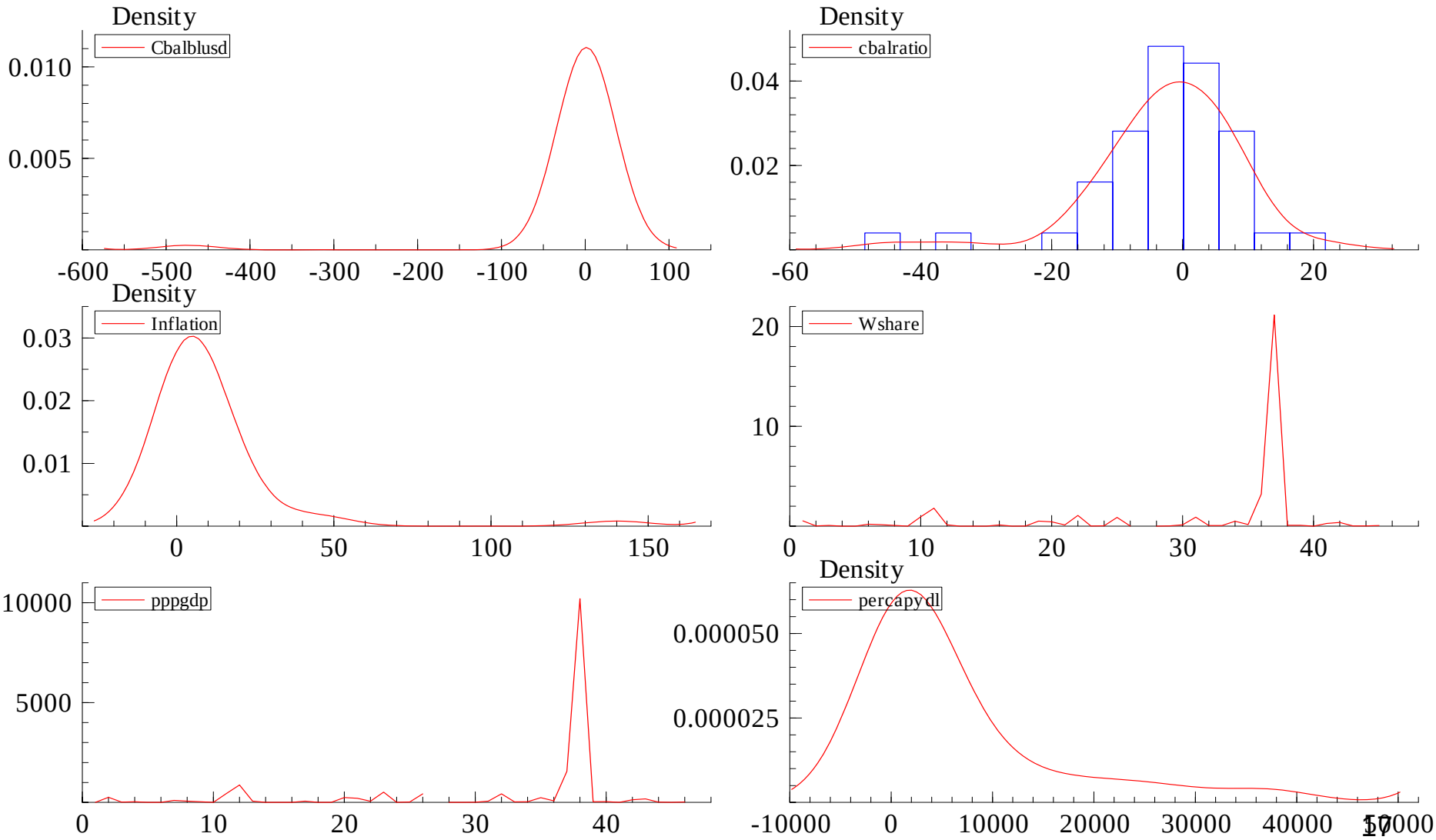
Forecasting Area, Production and Yield of Banana with a Simultaneous Equation Model



Trend Cycle and Irregular Components of Banana Production in Philippines (STAMP)



Annual Data does not have seasons


Density
Cbalblusd
0.010
0.005
-600
-500
-400
-300
-200
-100
0
100
Density
cbalratio
0.04
0.02
-60
-40
-20
0
20
Density
Inflation
0.03
0.02
0.01
0
50
100
150
Wshare
20
10
0
10
20
30
40
pppgdp
10000
5000
0
10
20
30
40
Density
percapydl
0.000050
0.000025
-10000
0
10000
20000
30000
40000
50000

References

- Greene WG(2000) Econometric Analysis, Prentice Hall.
- Stock JH and MW Watson (2003) Introduction to Econometrics, Pearson.
- Doornik J A and D.F. Hendry ((2003) PC-Give Volume I-III, GiveWin Timberlake Consultants Limited, London.
- Koopman SJ, AC Harvey, JA Doornik and N Shephard (1995) Structural Time Series Analyser, Modeller and Predictor, TCL.