Creating Global Database for Economic and Energy Modeling

Kyoto University

Second year doctoral course student

Shinichiro Fujimori

Background
## - What we are required -

- Global energy (GHG emission) modeling requires energy and economic data for calibration
  - Top-down model (multi-sector CGE) requires global Social Accounting Matrix (SAM)
  - Bottom-up model (End-use) requires energy demand data, economic activity data

- There are several database (statistics) relevant energy or economic data covering the world
  - Energy: IEA energy balance, enerdata
  - Economy: World Development Indicators (WB), OECD input-output tables, UNIDO industrial statistics, GTAP

- Most of the international statistics or database have some problems
  - Missing (time series or completely)
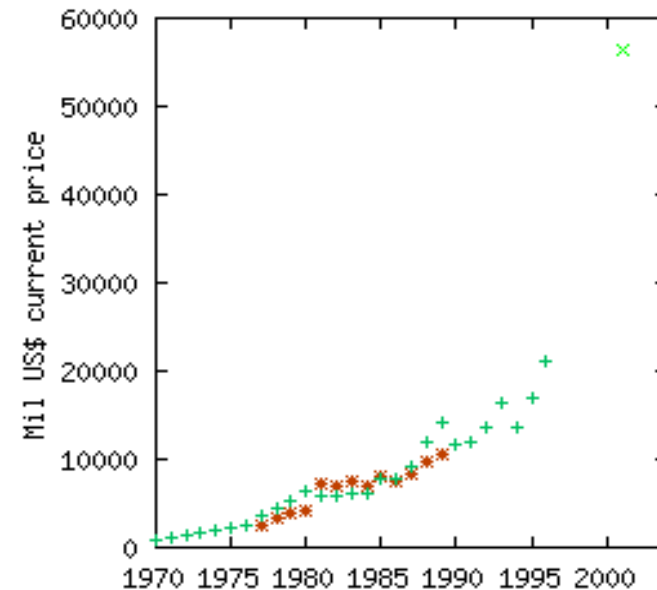 - Outliers
  - Inconsistency

Problems related to Statistics(1) -missing values-

Missing
Outliers
Inconsistency
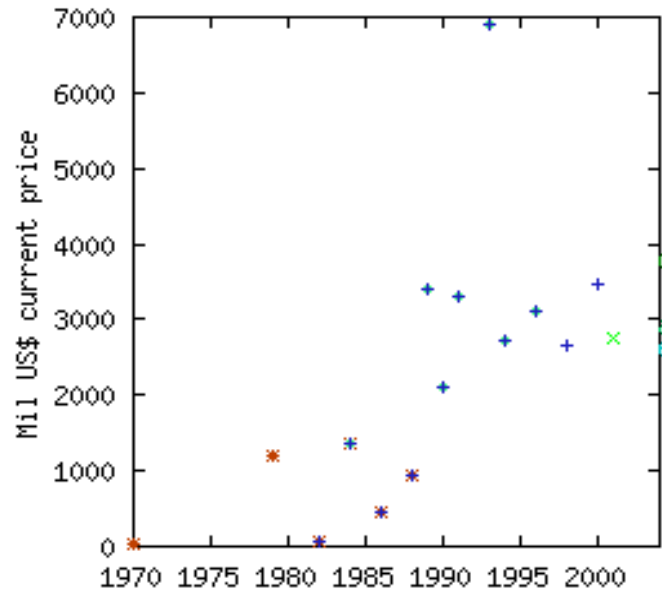
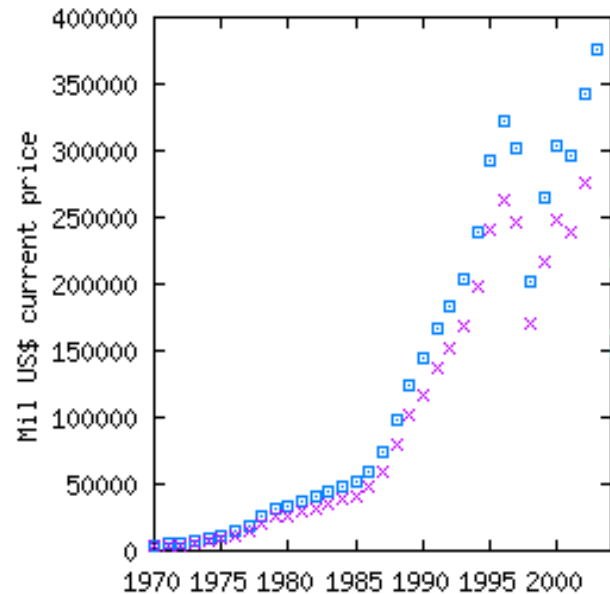
UNIDO
UN
GTAP
Takushoku
University

Thailand, Iron and steel

production (Mil US$)

China Paper and Pulp

production (Mil US$)

- Plotting the international statistic values
- ***Missing values*** are happened randomly or systematically
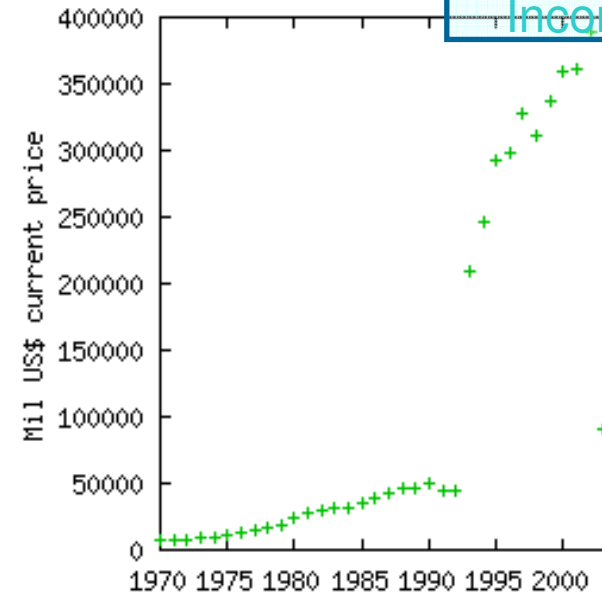- But how can we estimate these missing?

Problems related to Statistics(2)
-Outliers-

Missing
Outliers
Inconsistency

World bank
OECD
UN SNA

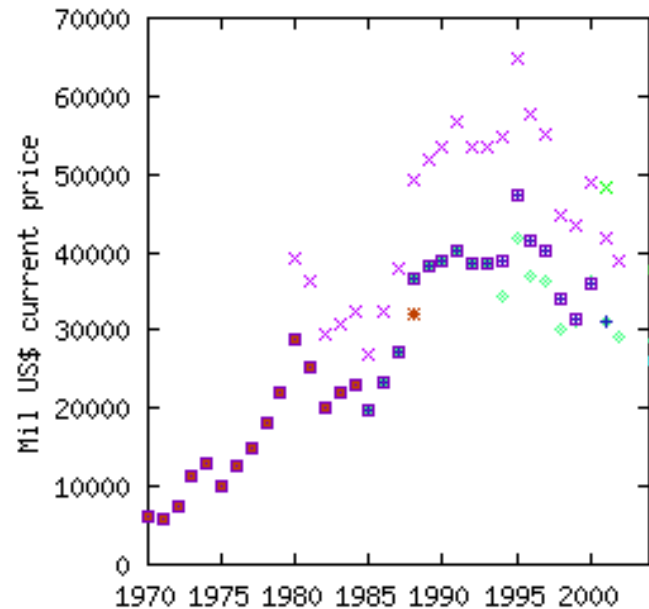Korea, Service sector
value added (Mil US$)

India, Industry
production (Mil US$)

- **_Outliers_**?
- Showing Real phenomenon?
- Data input mistakes?
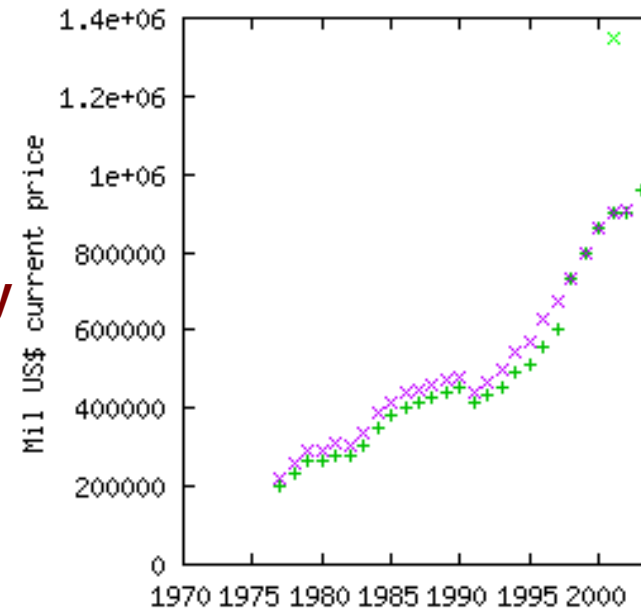- Classification change?

Problems related to Statistics(3)
-Inconsistency among statistics-

Outliers
Missing
Inconsistency

UNIDO
OECD
GTAP
UN SNA
UN industry

Japan Non-ferrous metal

production (Mil US$)

USA construction

production (Mil US$)

- **_Inconsistency among statistics_**
- Different sources?
- Classification differences?
- Problem with the estimation method?

Objectives

- Creating reliable and consistent database
  - Global trade balance
  - Regional Supply and Demand Balance
  - Without missing data
  - Consistent with reliable statistics
  - Linking with Monetary and Material or Energy

- Contributing to Global Modeling
  - Creating Global Database
    - Contents
      - Economy and Energy data
      - Material Flow and Stock (Energy and Basic materials)
      - Other Driving Forces (Population, labor forces and land use etc.)
    - Time series data
    - Regional detailed
    - Sector detailed

Framework of the Accounts (GAMMA)
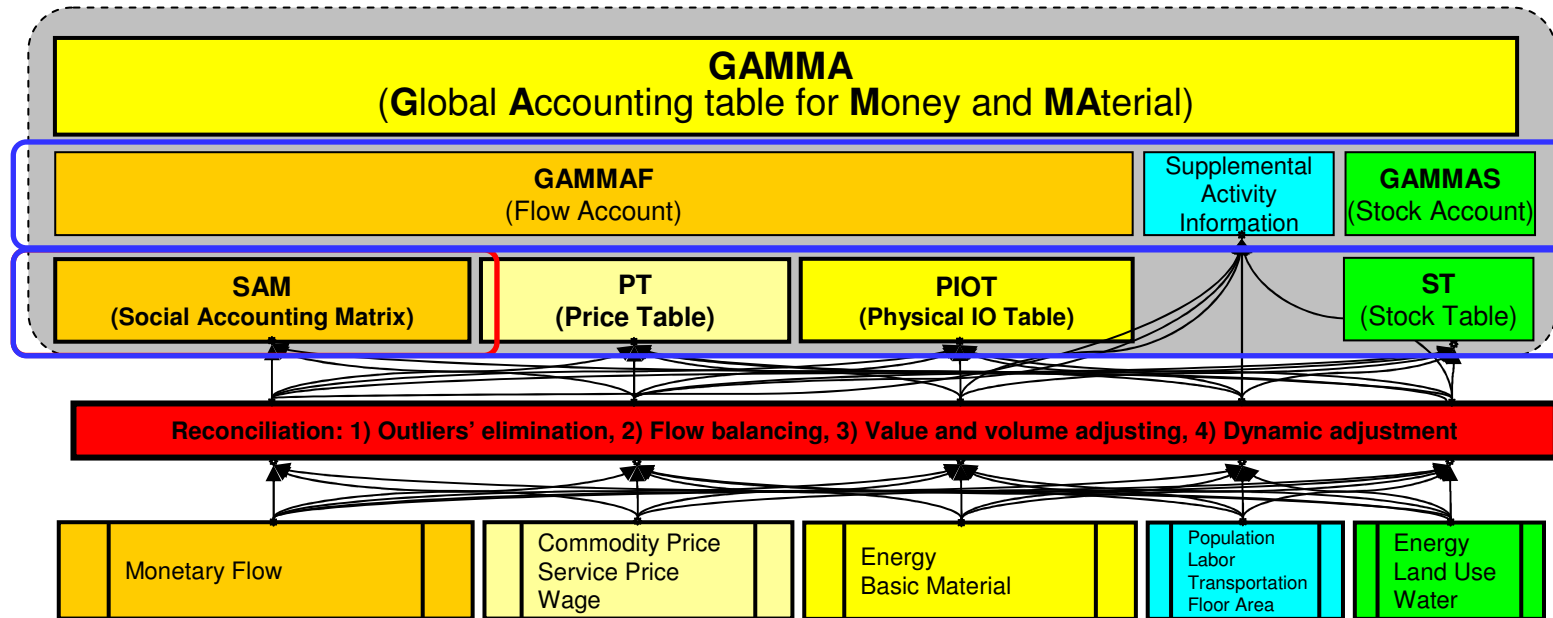
- **GAMMA** (**G**lobal **A**ccounting table for **M**oney and **MA**terial)
- 3 Kinds of Information
  - Flow, Stock and Supplemental activity information
- 4 Tables Describe flow and stocks
  - SAM (Social Accounting Matrix)
  - PT (Price Table),
 - PIOT (Physical IO Table),
  - ST (Stock Table)

Social Accounting Matrix(SAM)

	Commodity	Activity	Factor			Regional Household	Private Household	Government	Capital	Import Tax	Export Tax	Sales tax	Production Taxes	Direct Taxes	Transport Margin	Rest of the World	Trade Balance	Total
Commodity		Intermediate inputs					Household Consumption	Government Consumption	Capital Formation						Import Margin	Export		
Activity	Supply																	
Factor		Labor Inputs																
		Capital Inputs																
		Land Inputs																
Regional Household			Wage	Capital	Others					Import Tax	Export Tax	Indirect Tax	Production Tax	Income Tax				
Private Household						disposable income												
Government						Government Income												
Capital				Capital Depreciation		Saving											Trade Balance	
Import Tax	Import Tax																	
Export Tax	Export Tax																	
Sales tax		Indirect Tax						Indirect Tax										
Production Taxes		Production Tax																
Direct Taxes				Direct Tax														
Transport Margin	Import Margin																	
Rest of the World	Import																	
Total																		

- Describing one regional monetary flows
- Satisfying the balances of inputs and outputs
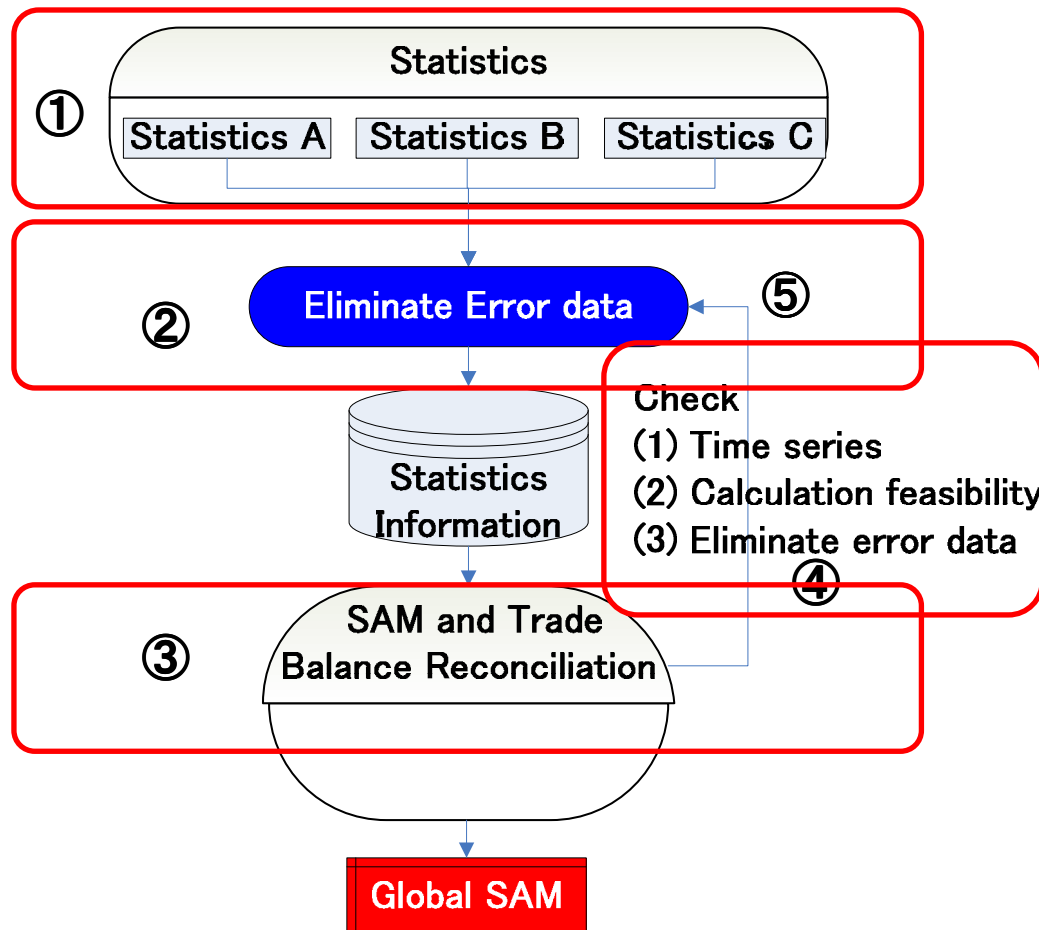
Methodology –Framework-
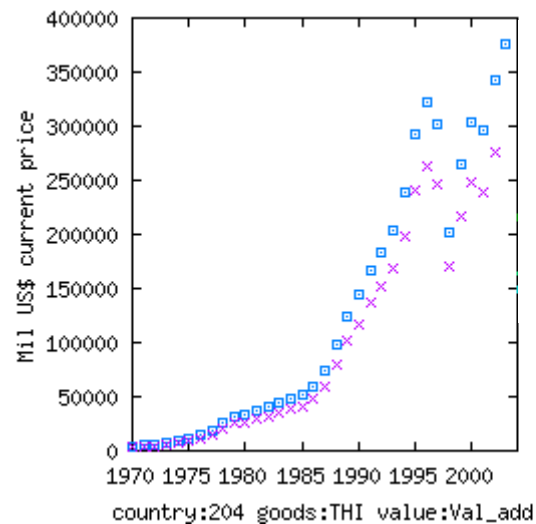

① Preparing various statistics
② Eliminate error data
   – Time series consistency
③ Reconciliation of SAM and trade balance
④ Re-check the time series and feasibility
   – Error data make Infeasibility
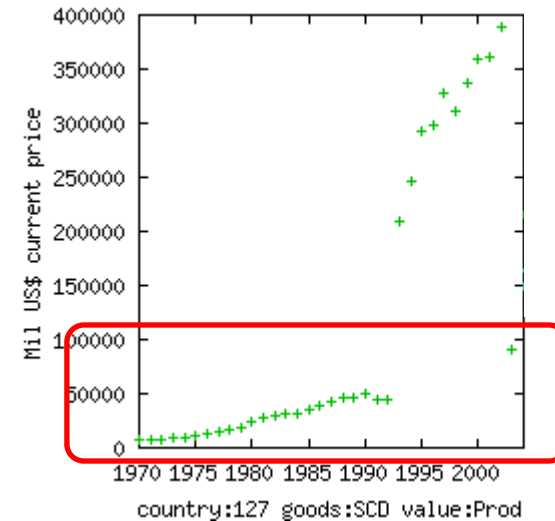   – Calculation feasibility
⑤ Eliminate error data again.
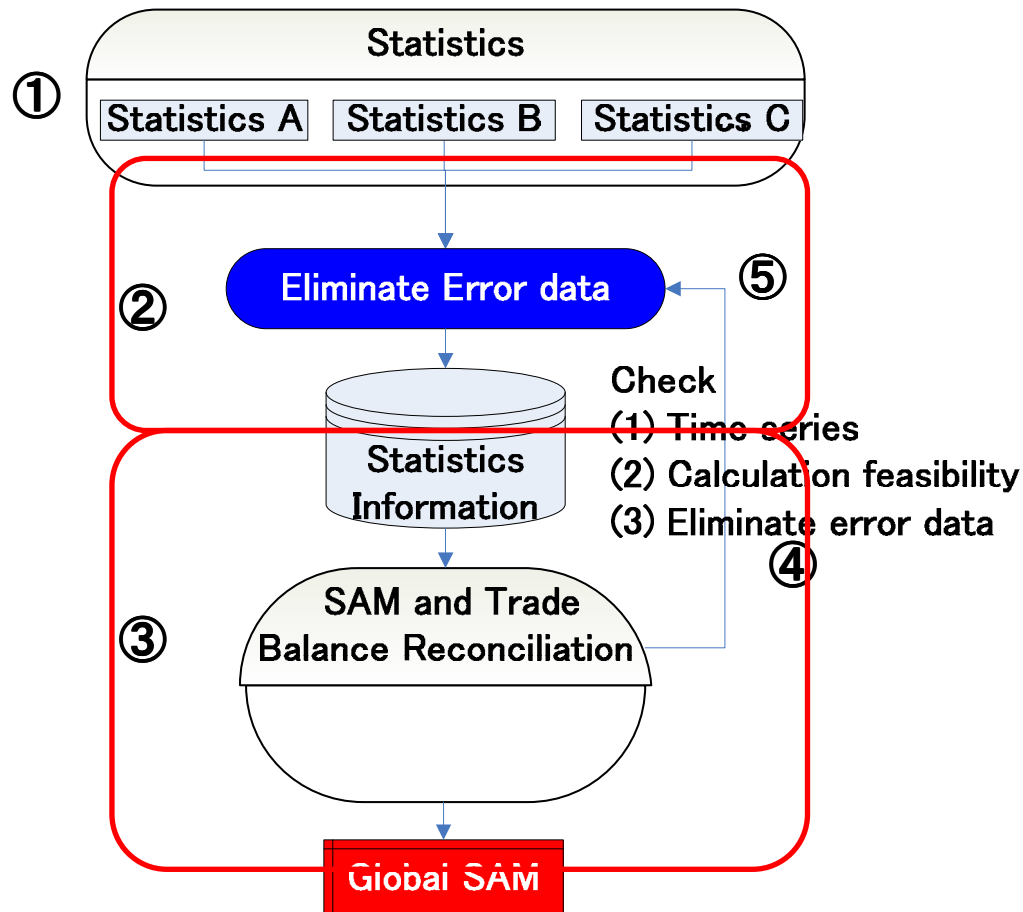
Problems related to Statistics(2)
-Outliers-

Korea, Service sector
value added (Mil US$)

World bank
OECD
**UN SNA**

India, Industry
production (Mil US$)

- At first we can not say they are wrong data
- In the calculation process,
  - The model compare with other industrial statistics
  - The model abort the solution because cannot satisfy the feasibility

Methodology –Framework-

① Preparing various statistics
② Eliminate error data
   – Time series consistency
   – Calculation feasibility.
   – Error data make Infeasibility
③ Reconciliation of SAM and trade balance
④ Re-check the time series and feasibility
⑤ Eliminate error data again.

Methodology –estimation procedure-
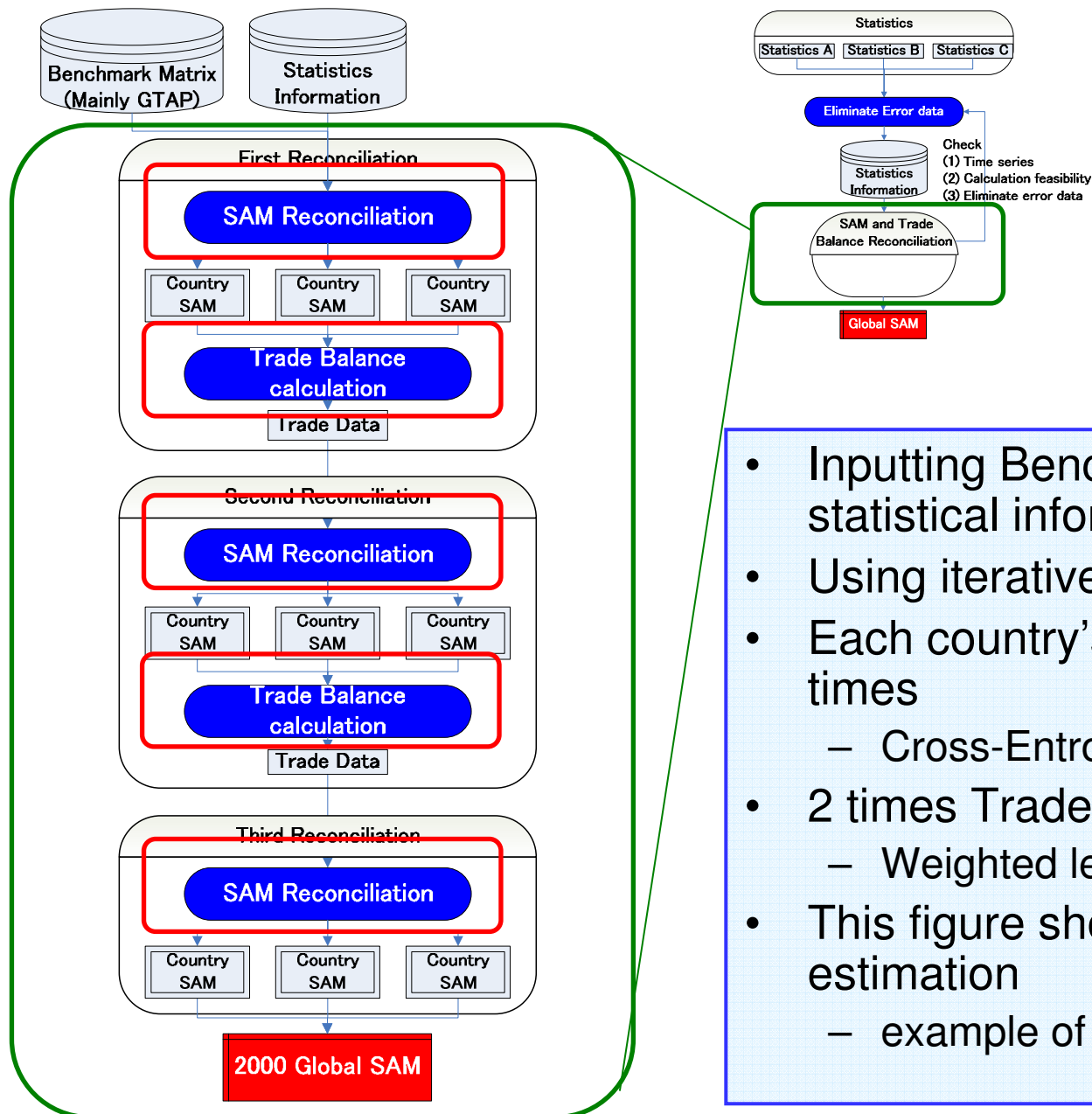
- Inputting Benchmark Matrix and other statistical information
- Using iterative procedure
- Each country's SAM are calculated 3 times
  - Cross-Entropy formulation
- 2 times Trade balance calculation
  - Weighted least squares method
- This figure shows one year matrix estimation
  - example of year 2000

Methodology
–step by step estimation procedure -
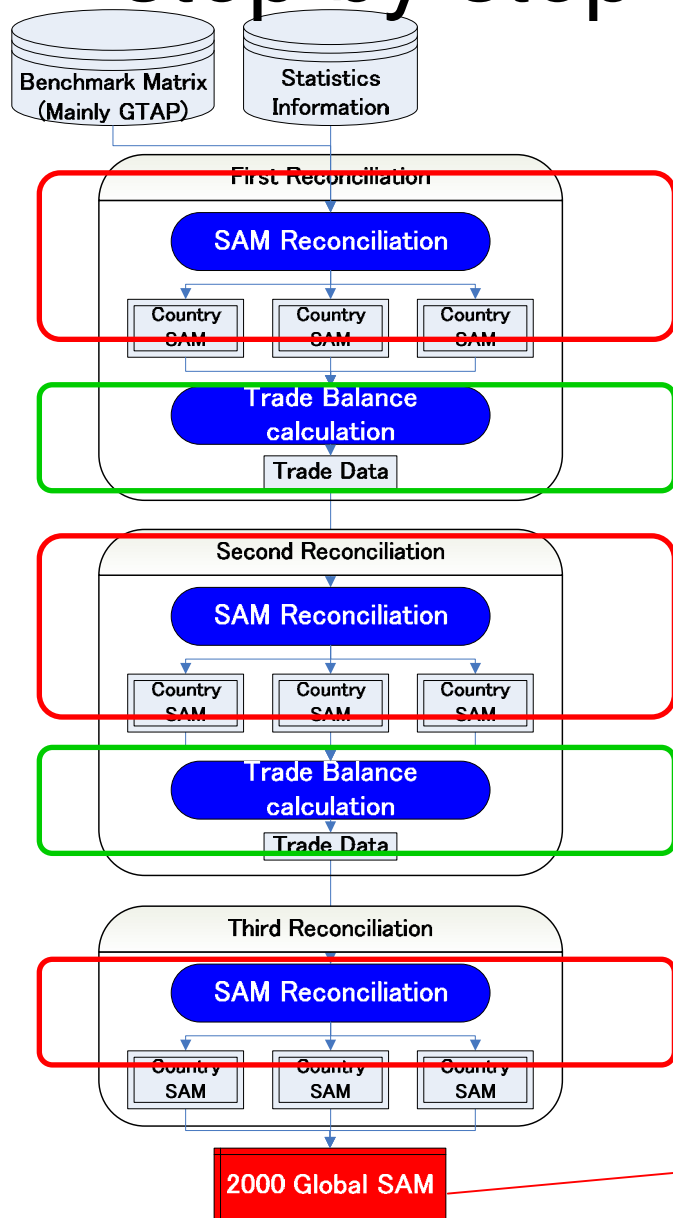
Each country's SAMs are estimated. But trade flows are not balanced in the world

Collect all country's trade data. Adjust the trade flow satisfying global trade balances.

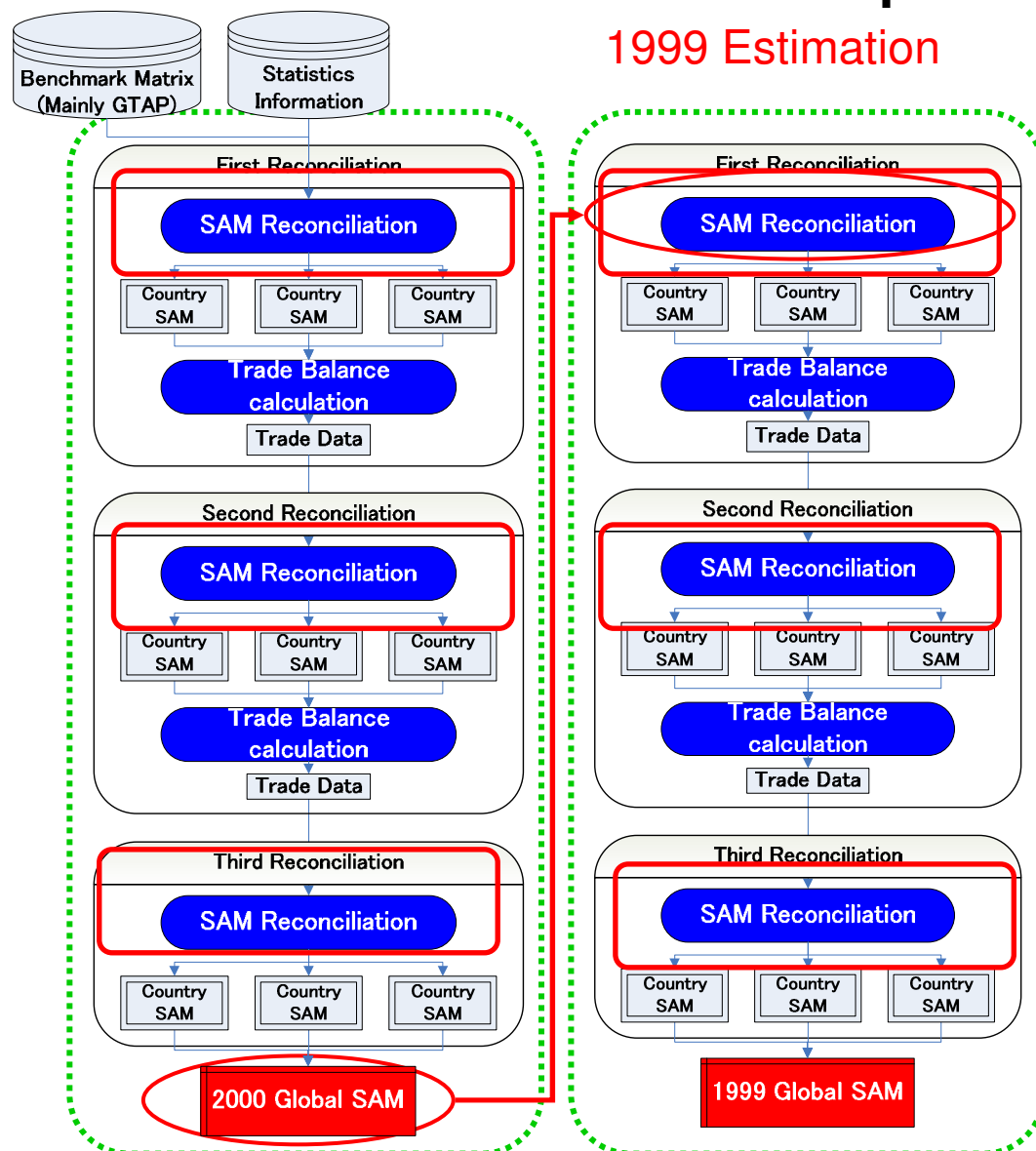**Second reconciliation uses previous reconciled SAM and balanced trade data**

Trade balance calculation same as first step

Third reconciliation fixes the balanced trade data

Get global SAM with satisfying trade balance

Methodology
–estimation procedure -

SAM Reconciliation (Basic formulation)

- ## Cross-entropy Method
  - Non-linear problem
  - Assuming input coefficient is similar to benchmark matrix

$$\min_{q_{i,j} > 0} \sum_i \sum_j p_{i,j} \ln \frac{p_{i,j}}{q_{i,j}}$$

→ $p_{i,j}$: **Input coefficient (Estimation)**

→ $q_{i,j}$: **Input coefficient (Benchmark)**

Where :

$$p_{i,j} = \frac{x_{i,j}}{\sum_i x_{i,j}} \quad \forall \, i, j$$

→ $x_{i,j}$: **SAM (Estimation)**

$$q_{i,j} = \frac{z_{i,j}}{\sum_i z_{i,j}} \quad \forall \, i, j$$

→ $z_{i,j}$: **Benchmark matrix**

SAM Reconciliation
(Additional economic information)

- Adding economic statistical information

- Assuming each statistics has errors

- Dealing with aggregated and disaggregated information

  - GDP, Value added

  - Commodity output, trade

  - Government consumption

  - Other information

$$\sum_{j \in J} \sum_{i \in I} g_{i,j}^{(k_1)} \cdot x_{i,j} = d_1^{(k_1)} \cdot e^{(k_1)} \qquad k_1 \in K_1$$

Statistical errors

Summation of the cells for the statistical information k_1

Statistical information

- Adding the penalty of the statistical errors in the objective function

$$\min \sum_i \sum_j p_{i,j} \ln \frac{p_{i,j}}{q_{i,j}} + \sum_{k1} F\left(e^{(k_1)}\right)$$

Penalty function of Statistics errors

Social Accounting Matrix(SAM)

	Commodity	Activity	Factor			Regional Household	Private Household	Government	Capital	Import Tax	Export Tax	Sales tax	Production Taxes	Direct Taxes	Transport Margin	Rest of the World	Trade Balance	Total
Commodity		Intermediate inputs					Household Consumption	Government Consumption	Capital Formation							Import Margin	Export	
Activity	Supply																	
Factor		Labor Inputs																
		Capital Inputs																
		Land Inputs																
Regional Household			Wage	Capital	Others					Import Tax	Export Tax	Indirect Tax	Production Tax	Income Tax				
Private Household						disposable income												
Government						Government Income												
Capital				Capital Depreciation		Saving											Trade Balance	
Import Tax	Import Tax																	
Export Tax	Export Tax																	
Sales tax		Indirect Tax						Indirect Tax										
Production Taxes		Production Tax																
Direct Taxes				Direct Tax														
Transport Margin	Import Margin																	
Rest of the World	Import																	
Total																		

Summation of the cells for the statistical information k_1

$$\sum_{j \in J} \sum_{i \in I} g_{i,j}^{(k_1)} \cdot x_{i,j} = d_1^{(k_1)} \cdot e^{(k_1)} \qquad k_1 \in K_1$$

Statistical errors

Statistical information

Total output of activity

Application

Used Statistics

- World 153 countries and regions
  - Including Integrated or breakup countries
- 1970 – 2003
- Commodity and Activity classification: 46
- Using more than 20 Statistics
- Tools
  - Solved by GAMS
  - Solver CONOPT3

	Statistics Name	Publisher
National Account	National Accounts Database	UN
General Statistics	World Development Indicators	World Bank (Mitchell, 2003)
	International Historical Statistics	
IO table	GTAP Database	GTAP
	OECD Input-Output Tables	OECD
	Asian International Input-Output Table	IDE
	Asean International Input-Output Table	IDE
Trade Statistics	Balance of Payments	IMF
	Commodity Trade Statistics Database	UN
	International Trade by Commodity statistics	OECD
	OECD Statistics on International Trade in Services	OECD
Industrial Statistics	General Industrial Statistics Database	UN
	Industrial Demand-Supply Balance Database at the 4-digit level of ISIC code	UNIDO
	Industrial Statistics Database at the 4-digit level of ISIC code	UNIDO
	Asian Long-term Statistics -Industrial Development-	Takushoku University
	FAOSTAT	FAO
	Structural Statistics for Industry and Services	OECD
	The OECD STAN database for Industrial Analysis	OECD
Country Statistics	Soviet Economic Statistical Series	
	The Soviet Economy 1970-1990 A Statistical Analysis	
	Taiwan National Account	
	HongKong's Merchandise Trade Statisitcs	

Results (Compare GDP with GTAP)

GDP comparison
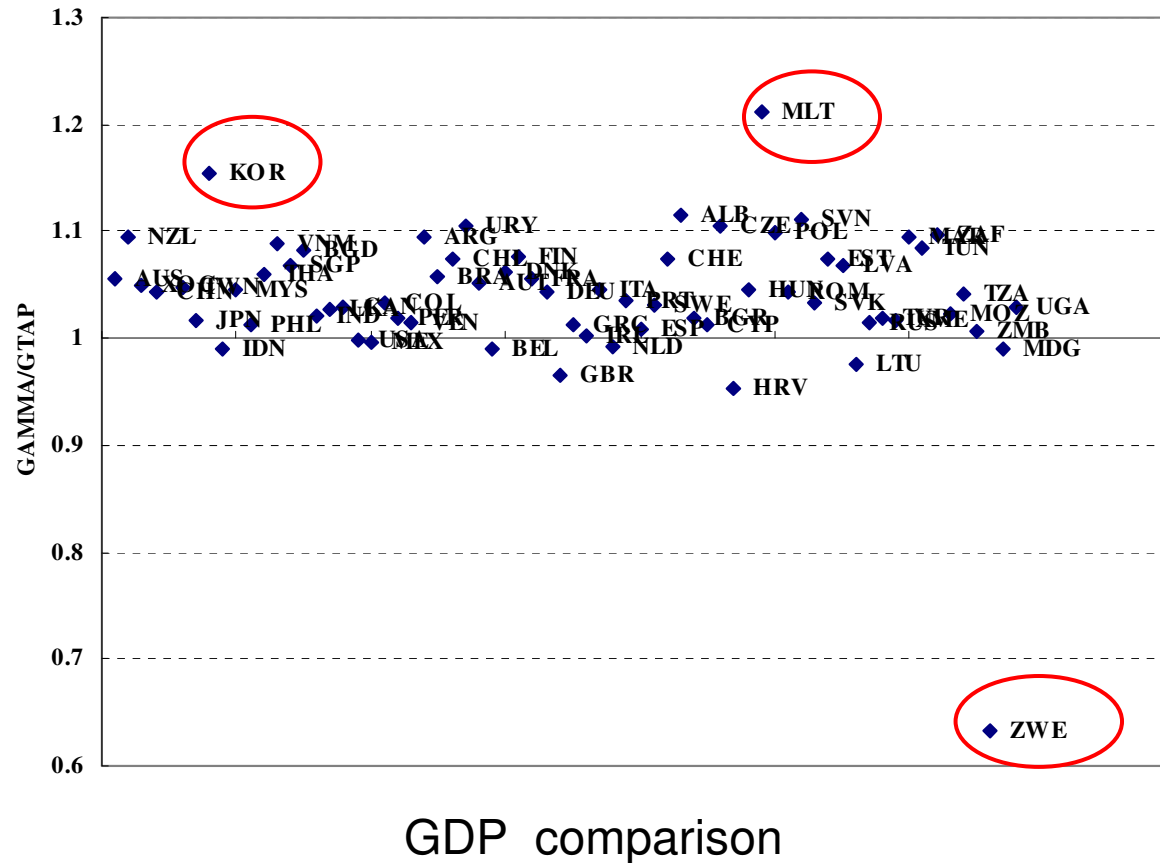
- GTAP is the benchmark matrix and it is used global CGE
- Plot Ratio of GAMMA to GTAP (Should be around 1)
- Some countries are different from GTAP.
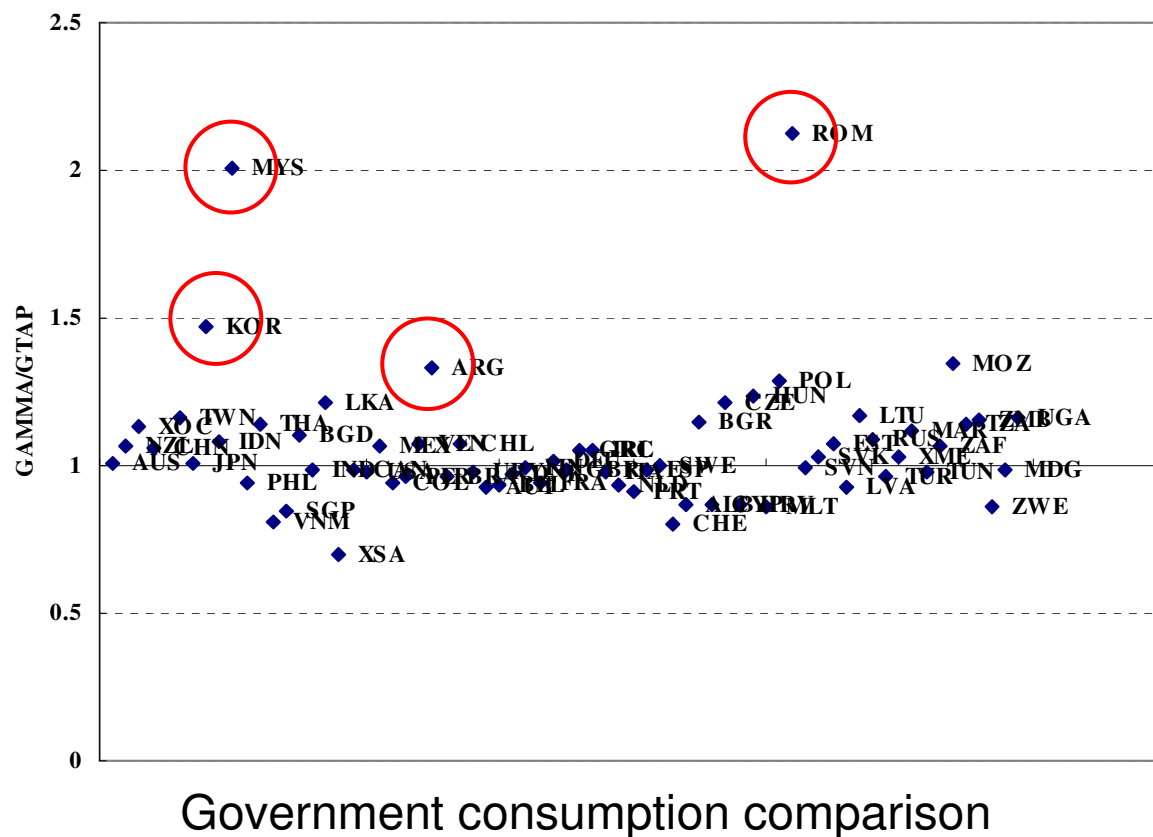
Results (Compare GDP with GTAP)
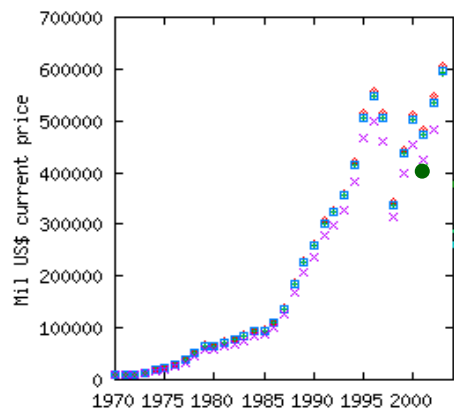
Korea
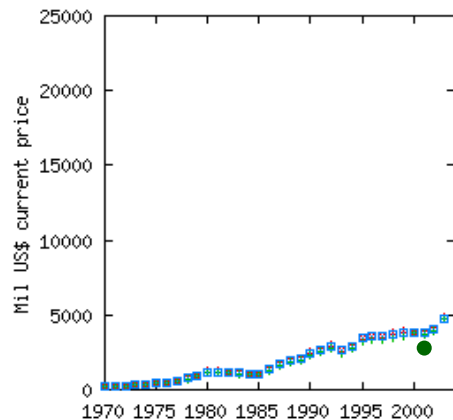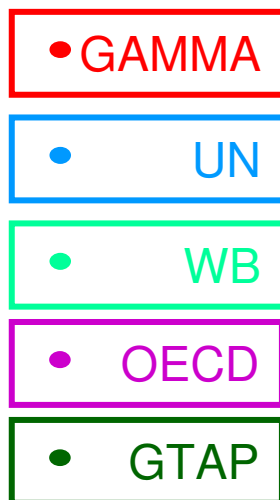
Malta

Zimbabwe
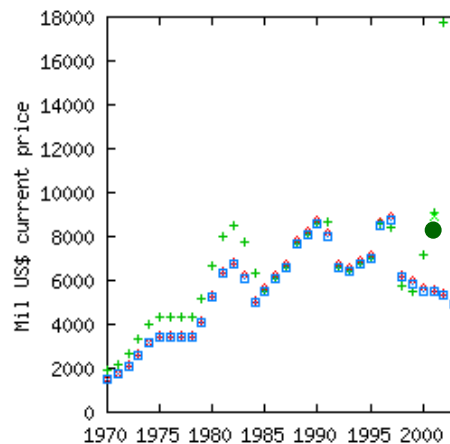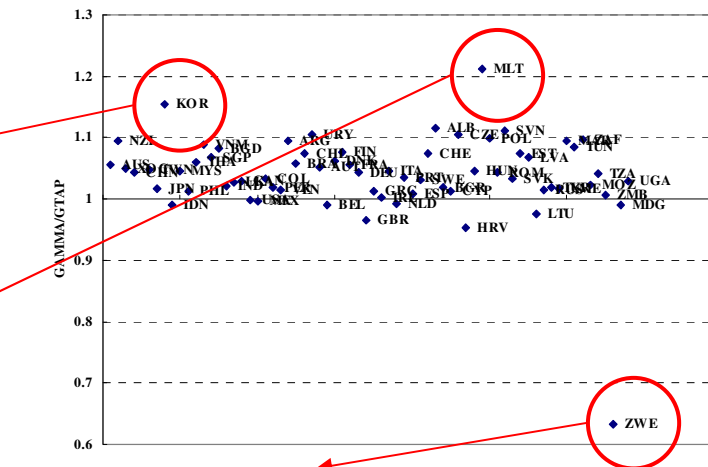
GAMMA
UN
WB
OECD
GTAP

- GAMMA is close to International statistics
- Differences are not so crucial??
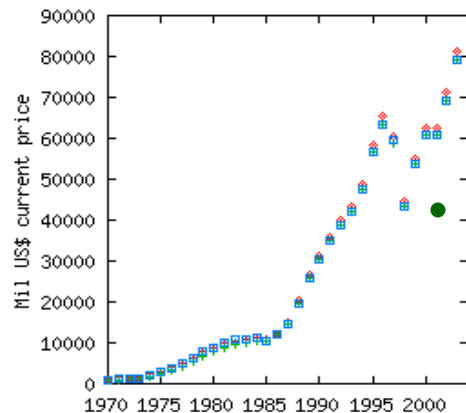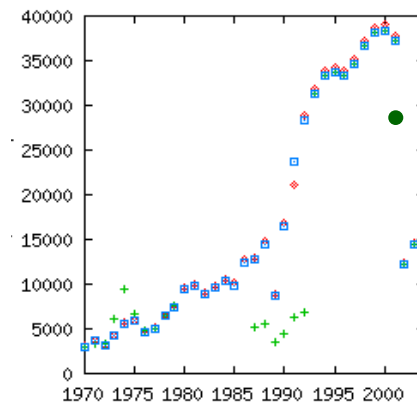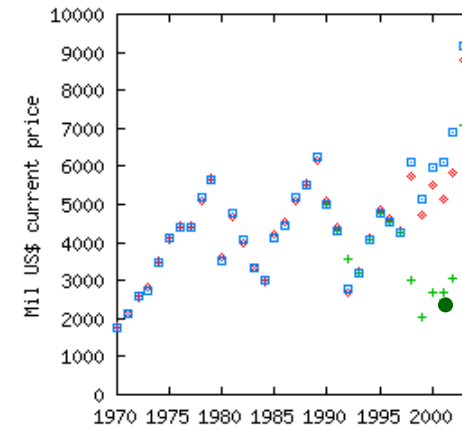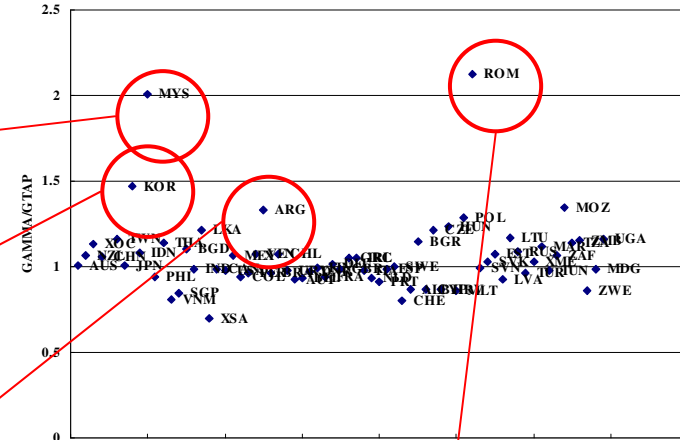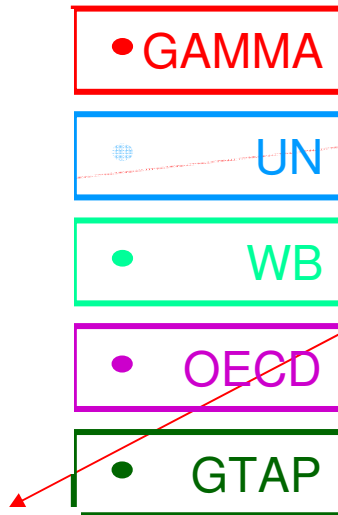
Results (Compare Government consumption with GTAP)

Government consumption comparison

- Plot Ratio of GAMMA to GTAP (Should be around 1)
- Some countries are different from GTAP.

Results (Compare Government consumption with GTAP)

Legend:
- GAMMA (red)
- UN (light blue)
- WB (green)
- OECD (magenta)
- GTAP (dark green)
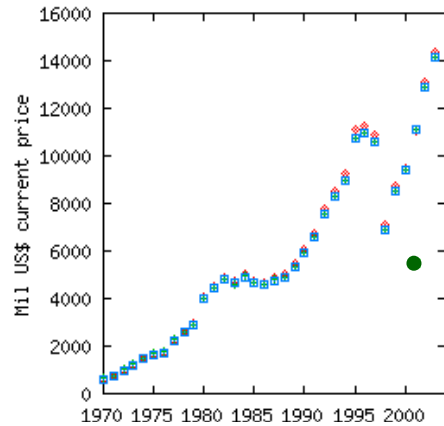
Malaysia

Korea

Argentina

Government consumption comparison

Romania

- Some countries show big differences from GTAP
- GAMMA is close to other international statistics

Problems related to Statistics(1)
-missing values-
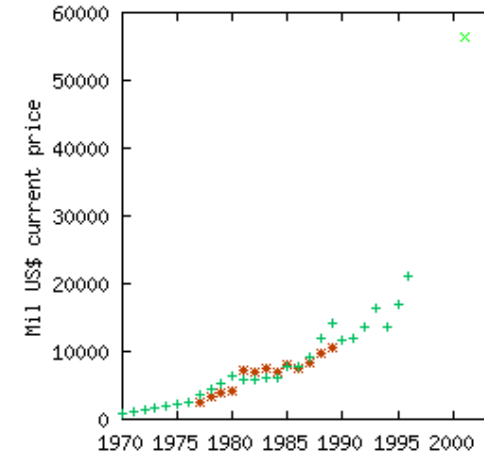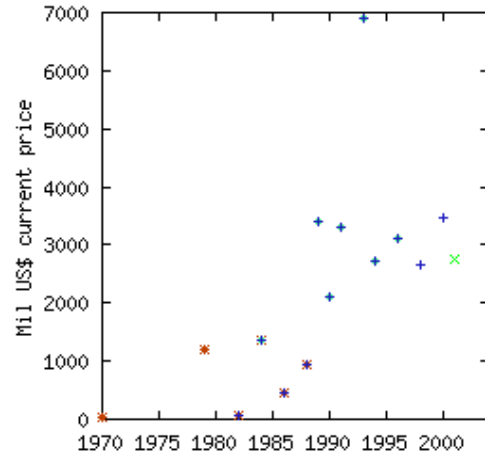
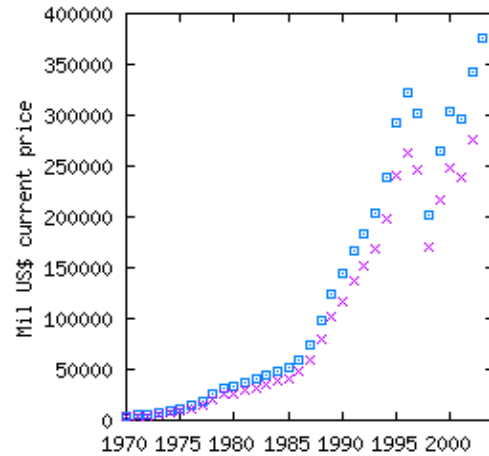

UNIDO
UN
GTAP
Takushoku
University
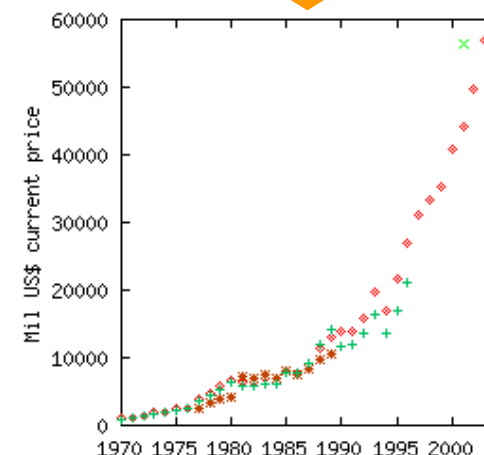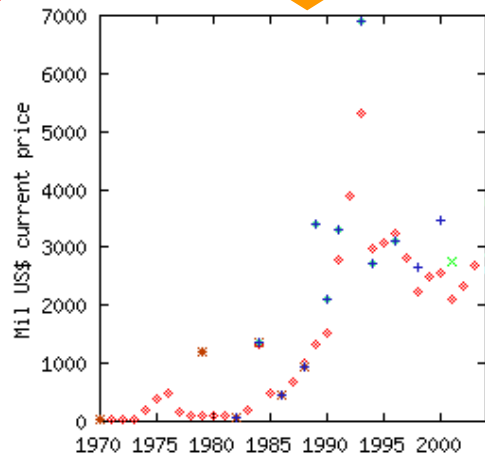
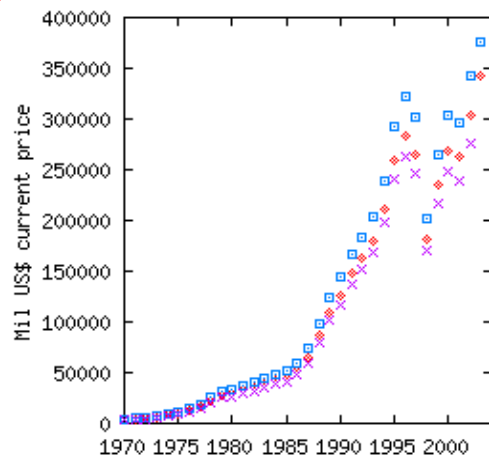**Estimation**

Thailand, Iron and steel
production (Mil US$)

China Paper and Pulp
production (Mil US$)
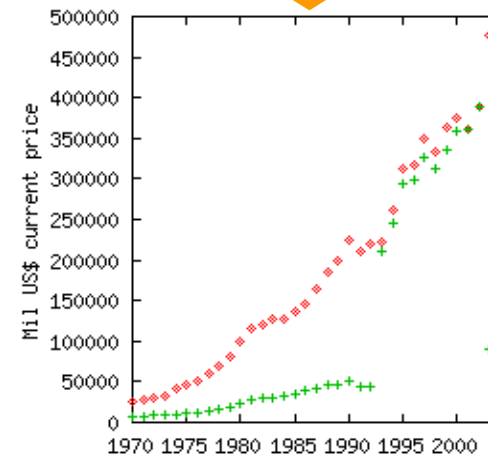
Problems related to Statistics(2)
-Outliers-
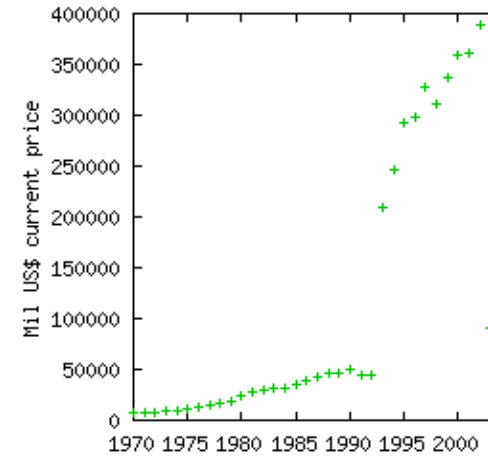
World bank
OECD
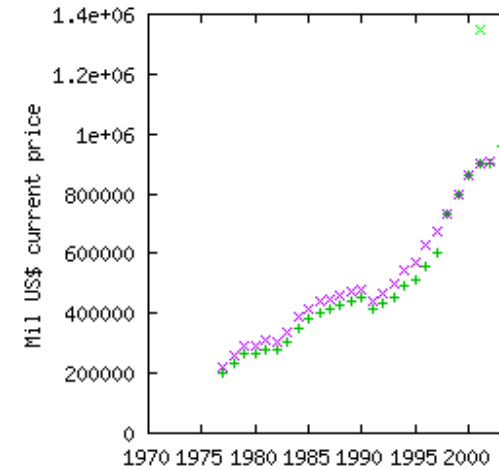**UN SNA**

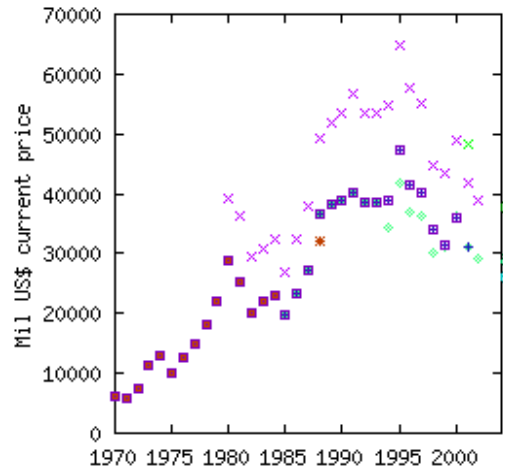Estimation

Korea, Service sector
value added (Mil US$)
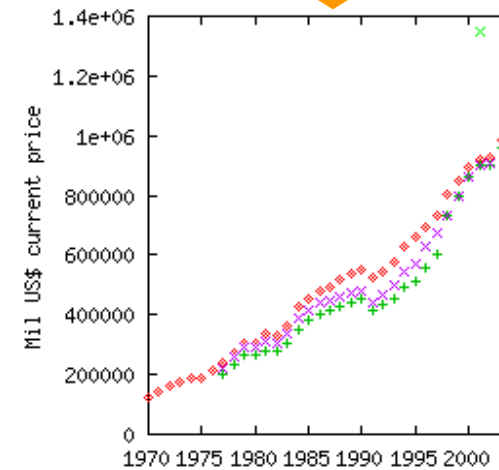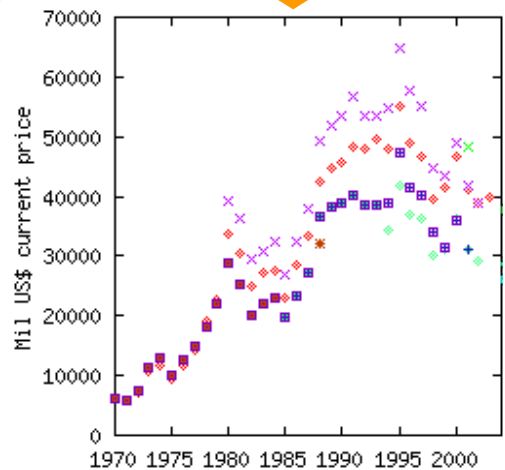
India, Industry
production (Mil US$)

Problems related to Statistics(3)
-Inconsistency between statistics-

UNIDO
OECD
GTAP
UN SNA
UN industry

Estimation

Japan Non-ferrous metal production (Mil US$)

USA construction production (Mil US$)

Final Remarks

- Propose framework of <u>GAMMA</u> database
  (Global Accounting table for Money and MAterial)

- Develop methodology for Creating Global SAM
  - Consistent and balanced
  - Regional, sector classification is detailed
    (153 countries and regions; 46 activities)
  - Time series data; 34 years

- Sometimes international statistics has errors, therefore such error data are eliminated

- Correspond with international statistics

Further Steps

- Filling material, price and stock account
- Analyze global material flow and stocks
- Modeling the material glow
  - considering international relationship
  - Location of energy-intensity industry