STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

FOOD AND AGRICULTURAL ORGANISATION (FAO)

Joint UNECE/EUROSTAT/FAO/OECD Meeting on Food and Agricultural Statistics in Europe
(Geneva, 2-4 July 2003)

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD)

FAOSTAT2 PROJECT and CountrySTAT**

Supporting paper submitted by FAO

---

* Paper posted on Internet as submitted by FAO.

** Prepared by Mr. E. Gillin, FAO.

FAOSTAT2 Project and CountrySTAT

The FAOSTAT system is one of FAO's most important corporate systems. It is a major component of FAO's information systems, contributing to the organization's strategic objective of collecting, analyzing, interpreting and disseminating information relating to nutrition, food and agriculture. It is at the core of FAO's World Agricultural Information Centre *(WAICENT)* through which access is given to FAO's vast store of information on agricultural and food topics – statistical data, documents, books, images, and maps.

FAOSTAT is a product that is well known throughout the UN, statistical and academic worlds. Policy formulators, decision makers and other stakeholders, both at the national and international levels are the primary users of the system. They regularly use FAOSTAT as a data source for analysis and decision making. Other users of the FAOSTAT system include FAO staff, the international community, researchers, private enterprise and the public at large. It is estimated that approximately one quarter of all visits to the FAO Web pages are made with the purpose of retrieving statistical data. FAOSTAT supports a subscriber base allowing users to perform bulk data downloads for analytical purposes. Data contained in the FAOSTAT system are regularly published both in hard copy yearbooks as well as on CDs.

The FAOSTAT working system has been operational for a decade. In recent years its technical and functional limitations have become more apparent, leading to growing user frustration. In late 2001, a proposal to proceed with a requirement analysis for the modernization of the FAOSTAT working system was endorsed and work is proceeding on the modernization of the FAOSTAT working system

The definition of requirements for a new FAOSTAT system commenced at that time. Interviews were conducted with representatives from all FAO organizational units that are currently providers of data to FAOSTAT as well as those who compile data that potentially could be included in a new FAOSTAT system. Principal internal users of FAOSTAT data were also contacted for their views on needed improvements to the system. A small representative sample of external users of the system was also queried.

Objectives of the New FAOSTAT2 System

The anticipated end result of the FAOSTAT2 initiative is a new system that will meet FAO 's Strategic Framework Objective of *"an integrated information resource base with current, relevant and reliable statistics, information and knowledge made accessible to all FAO clients."* To achieve that goal, FAOSTAT2 must specifically address the following objectives:

- Provide an improved user interface, streamlined system processes, and a stable and reliable technical environment for the FAOSTAT working system.
- Improve the quality of the data by providing robust tools for compiling, validating, estimating and analyzing data both at FAO headquarters as well as at the country level.
- Incorporate new user requirements for system functionality and access to new data sets.
- Improve user access to FAOSTAT data by enhancing and creating new mechanisms for data dissemination, including access to data across domains.
- Enhance data integrity by ensuring that appropriate methodologies and data standards are consistently applied.

A stable and reliable technical architecture

The proposed project approach recommends the use of components that have been created for similar projects already underway within FAO as a "quick start" method for building critical system functionality. Initial project outputs would address the more pressing technical problems and major user requirements. Major improvements in system stability and reliability should be realized early in the project.

Incorporate New User Requirements

A number of new user requirements would be incorporated into the new FAOSTAT2. These include improved or new system functionality, as well as expanded data requirements.

Improved or New System Functionality: New functionality would focus primarily on tools to facilitate data capture and improve data quality: data exchange standards; flexible, yet comprehensive, data editing routines; data estimation support and tools to be used at country-level. Software would be developed to enable users to easily extract and analyze FAOSTAT2 data.

*Expanded Data Requirements***:** the requirement analysis revealed a need for FAOSTAT2 to accommodate diverse types of data, which go beyond the standard cross section and annual time series currently maintained within the system. Many data sets currently maintained outside of the FAOSTAT system could be incorporated into an extended FAOSTAT data model. To accommodate these requirements, the FAOSTAT2 system should be capable of supporting the following:

- *Data of varying periodicity:* FAOSTAT2 would support data capture at any frequency and interval, in addition to the standard annual time series data. This would include data captured monthly, on a market year basis, etc., as well as data captured on a less frequent basis.
- *Data projections: FAO* projections would be included for any category of data held in the system. This would provide a more complete picture of anticipated trends based on historical information.
- *Sub-national data:* The new system would provide the capacity to store and report on country data that is captured at the sub-national or administrative unit level.
- *Geo-referenced data: FAO* supports a number of spatial information systems that contain a wealth of information on geographic conditions impacting food and agriculture production. Integration of FAOSTAT statistical data with the various FAO spatial information systems would enhance FAO's analytical capacity in the food and agriculture arena. To support this requirement, the FAOSTAT2 system would provide the capability to attach a "geo-reference" to any data maintained within the system.
- *Data from external sources:* Data available from external sources is often useful and necessary to assist FAO staff in their analyses. A flexible and easy-to-use mechanism is required to access data from such sources as the IMF and World Bank and to merge it with FAOSTAT statistical data.

- *Meta data:* Meta data enhances understanding of any given data item within the system by documenting its definition, history of its values, methodology used in its collection, national contacts, etc. This information is useful to statisticians who compile, validate and analyze the data, as well as to the users, both internal and external to FAO, who access the data. The construction of a meta data repository and its integration with FAOSTAT statistical data is an essential component of the new FAOSTAT2 system.
- *Reference data*: Data used in supporting calculations and validations, e.g. conversion factors, country codes, editing rules, etc., would be available in and accessible to users of FAOSTAT2.

The Faostat2 Data Flow is shown in Figure 1.

Figure 1. FAOSTAT2 Data Flow
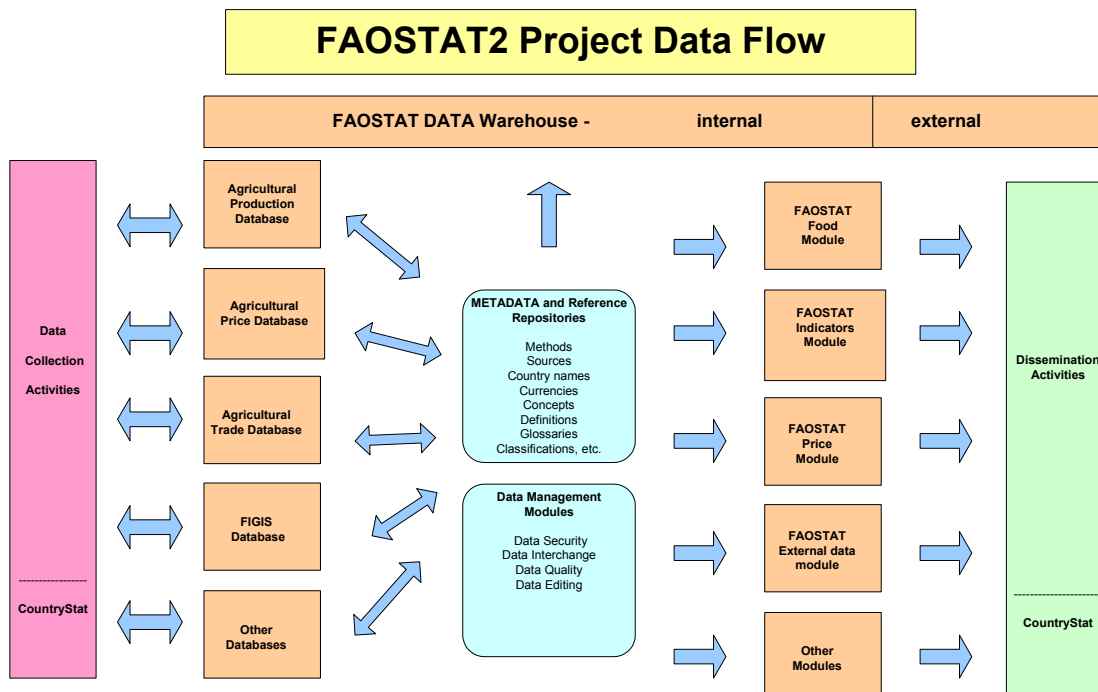
It is proposed that the new FAOSTAT system revolves around a core FAOSTAT module (see figure 2.) with distributed database modules around the core module. This model provides a flexible approach as the satellite databases need only to have linkages via correspondence tables to the core and other modules to enable data interchange. The core module will have standard statistical metadata elements to facilitate data interchange with the other database modules. Only selected statistical data will be included in the core module, such as the core food and price modules (see figures 3 and 4.).

Figure 2.   FAOSTAT2 Core and sample draft satellite database system
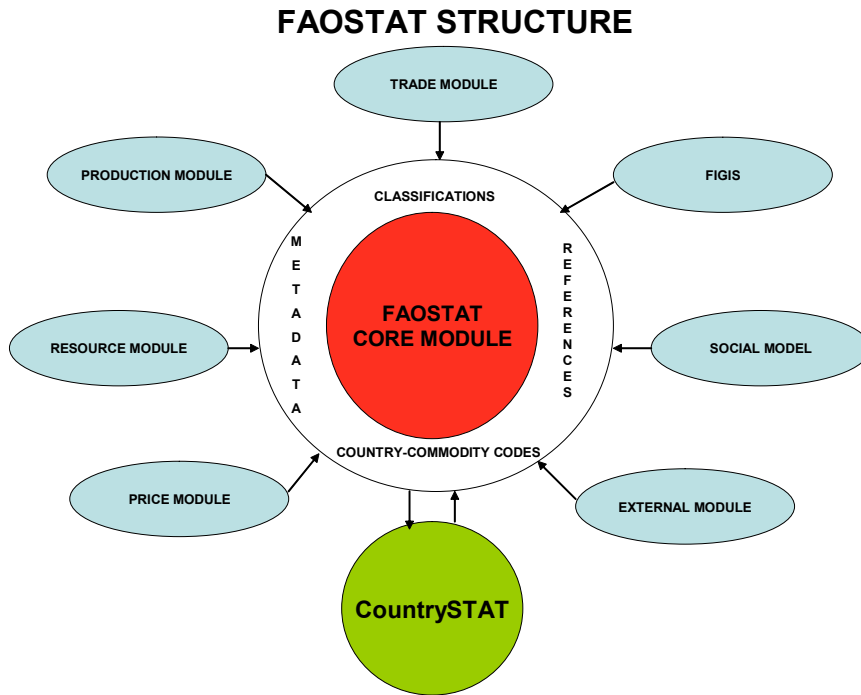
FAOSTAT STRUCTURE

Figure 3.   FAOSTAT2 Food Module

FAOSTAT FOOD MODULE

Figure 4.   Example of a Satellite module

FAOSTAT PRICE MODULE

	UNIT PRICES (National Currency)	**INDEX NUMBERS** Base =100

OUTPUTS

INPUTS

PRODUCER PRICE
WHOLESALE PRICE
SUPPORT PRICE
RETAIL PRICE
EXPORT PRICE
IMPORT PRICE
WORLD PRICE
OTHER PRICES

Producer Price Index
Wholesale Price Index
Consumer Price Index
Food Price Index
Export Price Index
Import Price Index
Official Exchange Rate
Market Exchange Rate

AGREGATE INDEXES OF PRICES PAID AND RECIVED AND EXCHANGE RATES
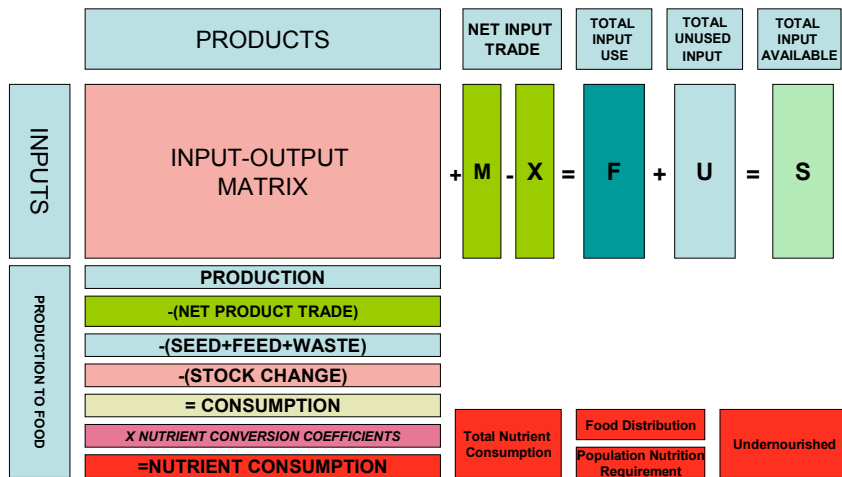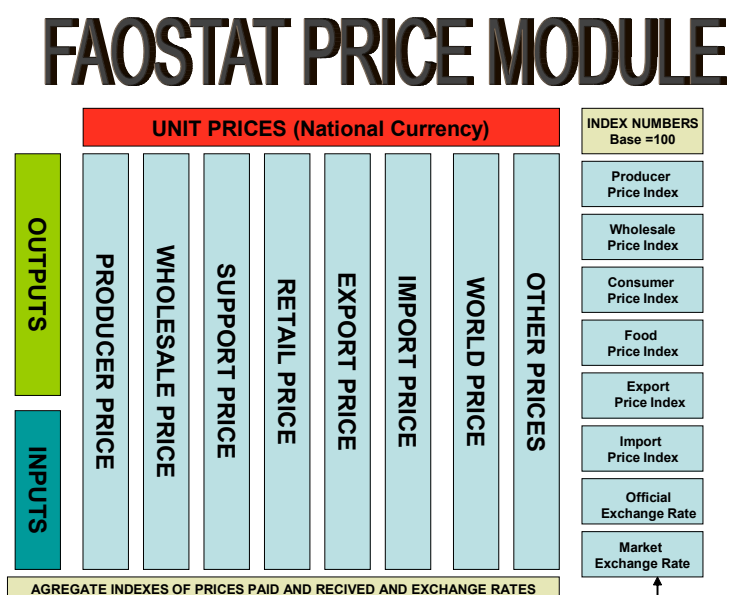
Improve data quality

Several mechanisms for ensuring data quality (in terms of accuracy and completeness) are envisioned for the FAOSTAT2 system. All data entered into the system would be submitted to a series of rigorous editing and consistency checks. System editing would be performed in real-time to identify data errors early in the process and increase data entry efficiency.

Where a country has not provided complete times series data, the system would provide appropriate tools for estimating missing values. These tools would include interpolation and extrapolation algorithms as well as statistical models developed using analytical software packages.

The development and publication of an XML-based standard for the exchange of food and agricultural statistics would also facilitate improvements in the quality of FAOSTAT2 data. This standard would be distributed to member countries and other external partners who would be requested to provide their data in files formatted according to the published standard. Data provided in this fashion would be loaded, edited and validated, thus saving FAO staff time and improving data quality.

FAOSTAT data quality ultimately depends on the quality of the data reported by member countries. This could be facilitated if countries entered and processed data online, which is currently possible, to a limited extent, through the Web-based Virtual Questionnaire. While this method would continue to be an option under the new FAOSTAT2 system, another option would also be made available: CountrySTAT, a scaled-down version of the FAOSTAT2 application, which would provide countries with functionality to compile, validate, analyze and disseminate their national data. Outputs from CountrySTAT could then be easily loaded into FAOSTAT2 for further dissemination through FAO's web site, publications and CD-Roms.

While countries would be encouraged to implement CountrySTAT, it is understood that all existing methods for data capture and transmittal - paper questionnaires, Virtual Questionnaire, electronic files - would continue to exist under the new FAOSTAT2 system.

CountrySTAT

CountrySTAT should be considered as a separate project to be developed, in parallel to FAOSTAT2. CountrySTAT should be conceived as a 'mirror' image of FAOSTAT2 to take maximum advantage of the development work that would be done anyway for FAOSTAT2. There was already a demand for information systems at the country level (Democratic Republic of Congo, Ethiopia, Lebanon, Syria, Vietnam) which was expected to increase rapidly. At present FAO is having to formulate one-off projects to accommodate this need which were proving to be expensive and resulting in tailor made systems at the country level. Considerable advantage was seen in having an 'off the shelf' system that could be put into countries.

In terms of the **scope of CountrySTAT,** that it should provide a storage, verification, validation, analysis and dissemination system appropriate to a system of food and agricultural statistics at the national level and should thus have the capacity to store sub-national data and geo-referenced data as had already been agreed would be part of FAOSTAT2. It should also have the basic tools that will be developed for FAOSTAT2 to verify, validate and derive the data (food balance sheets/supply utilization accounts) so that increased responsibility and ownership could be taken by the countries for the quality and consistency of their data. No conflict was seen between **CountrySTAT and FAO's role in independently verifying reported data**. By improving the quality of the data at the country level, FAO's work should be reduced as there would be fewer problems to sort out. By giving the country's greater responsibility for their own data and the tools to explore the quality of the data, long term improvements in the quality of the data provided by countries could be expected.

In terms of the **relationship between FAOSTAT2 and CountrySTAT** it is felt that the proposal that CountrySTAT should be a 'mirror' image of FAOSTAT2 meant that the relationship between the two would be extremely close. While the functionality would remain the same, the main differences would be the ability for CountrySTAT to operate in a scaled down version with no licensing costs or additional software, hardware requirements for the Country. CountrySTAT would be able to integrate within different IT environments of developing countries and be able to be easily updated, modified, maintained and sustained; taking into consideration what could be limited country resources and skills.

On the last issue, **how to ensure consistency of data,** by defining CountrySTAT as a country version of FAOSTAT2 this could only improve the consistency of data since the countries would be adopting an identical approach to data verification and validation as that used by FAO.

Given that the development of such a system is seen primarily as a by-product of the development of FAOSTAT2, it makes sense to develop the two in tandem and within the FAOSTAT2 project. The approach would provide a valuable statistical compilation validation, generation and dissemination tool that would provide valuable Country outputs such as Food Balance Sheets and be feasible within the objectives and scope of the FAOSTAT2 project.

New mechanisms for data dissemination

The original proposal for the modernisation of FAOSTAT focused solely on the working system, i.e., the underlying system used to compile, validate and analyze statistical data. The working system clearly has the most pressing problems and is in need of replacement. However, the analysis of system requirements revealed the need to also address the redevelopment of the dissemination system.

The dissemination system is the visible portion of the FAOSTAT, the component that most external users are familiar with. It is through the dissemination system that data can be accessed for further use. Enhancing the functionality of the dissemination system is seen as a means for strengthening the capacity of internal and external users of statistical data to perform more substantive analytical work.

Several opportunities exist to enhance and extend the functionality of the dissemination system. An improved interface, with the capability to visualize FAOSTAT2 data through graphs and maps, would enable users to highlight trends, anomalies and areas of concern. Key country indicators would also provide a useful vehicle to assess a situation at a glance.

Data dissemination mechanisms would be enhanced to support the automatic generation of high quality system outputs. New alternatives for publishing FAOSTAT2 data, such as dynamic online reporting, would be introduced. Dissemination monitoring tools would be expanded to enable FAO to analyze patterns of system usage and to better tailor its services to meet the needs of its users. Initial work has started on the new Internet user interface to FAOSTAT2. Examples of the new user interface are shown in Figures 5, 6 and 7 below.

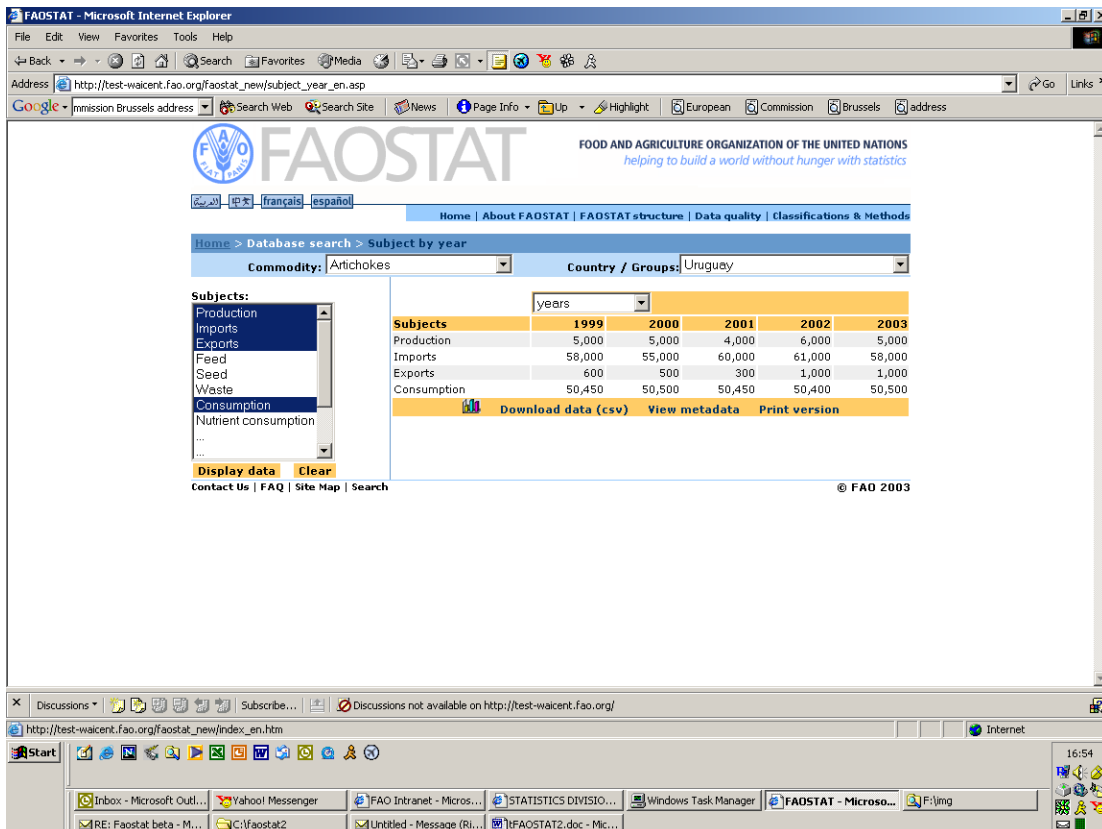**Figure 5.   FAOSTAT2 – mock-up of new homepage.**
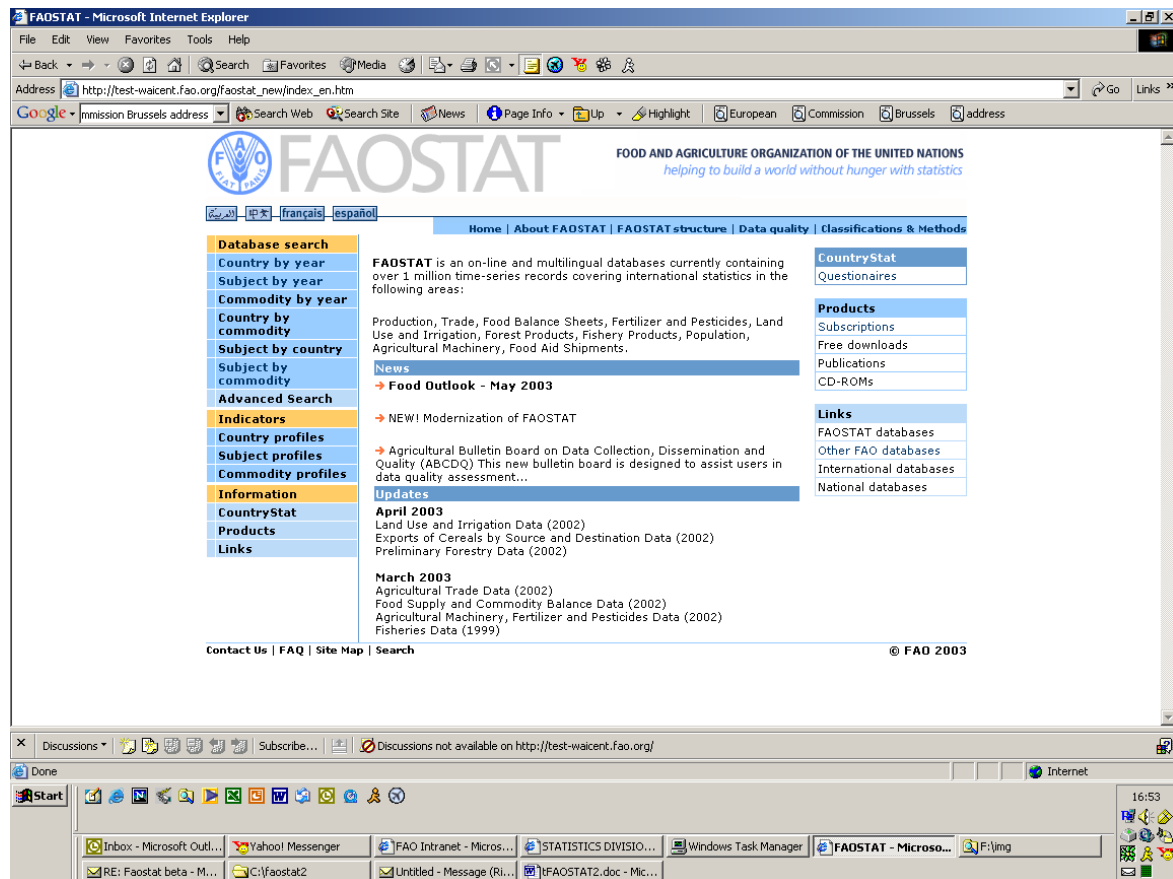
Figure 6. **FAOSTAT2 - mock-up of data selection of artichoke production by country by years.**
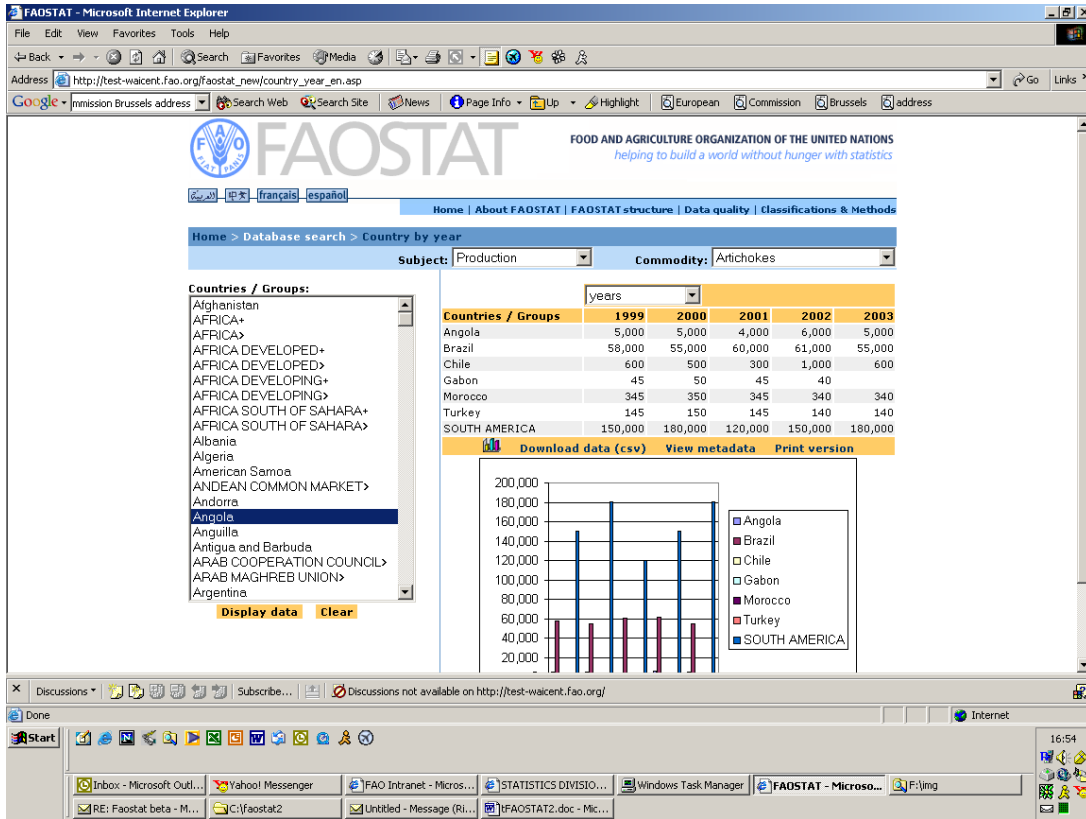
Figure 7. **FAOSTAT2 - mock-up of data selection of production, imports, exports and consumption of artichokes for Uruguay for various years.**

Appropriate methodologies and data standards

As the scope of the FAOSTAT system expands to include data from more diverse sources and to address the needs of more organizational units within FAO, the need to expand and reconfirm common data concepts, definitions, codes and methodologies becomes increasingly apparent. The FAOSTAT2 project offers an excellent opportunity to revisit existing methodologies and to identify areas for improvement, simplification and standardization. The requirement analysis recognized the need to address data management and statistical methodology issues. This includes identifying the appropriate organizational mechanisms to support ongoing data management and statistical methodology review. A working group has been established to define the terms of reference for this work, which should be undertaken prior to the commencement of the FAOSTAT2 project.

Proposed Project Approach

FAOSTAT2 would be implemented using a phased approach. The project would be divided into three sub-projects (working system, dissemination system and integrated corporate database). Work on these sub-projects would proceed relatively independently of each other.

Initial outputs of the project would focus on improvements to the current system and resolution of the most pressing problems. New system functionality would be capable of running on the current FAOSTAT system, thus enabling users to gain early experience with individual system components, before the complete, integrated system is delivered.

The FAOSTAT2 project would attempt to incorporate or use software components that have already been developed within the Organization. Rapid application development and prototyping would be employed, so that users can react to the proposed system and quickly provide feedback to developers.

The FAOSTAT2 development effort would be user-driven, involving key FAO staff at every level of the project. Data producers and data users would participate in the project from its inception, working as part of the project team to design and test the new system.

This approach to the development of FAOSTAT2 should provide numerous benefits:

- Users will see tangible progress, which should build goodwill and enthusiasm for the final system.
- Incremental introduction of system functionality should facilitate acceptance of the new system.
- User involvement in the project from its inception should ensure that users' needs are met and that there are no "surprises" once the full system is implemented.
- The FAOSTAT2 system will evolve as users become familiar with individual system components and make more informed suggestions for changes to the final system. This should improve the system design and create a sense of ownership.
- Using already-developed software components should save time and effort.