

Scientific final Report- Riccardo Da Re

Cost Action: FP 1203 “European Non-Wood Forest Products Network”

STSM title: International trade network analysis.

Reference: COST-STSM-ECOST-STSM-FP1203-010214-039672

STSM dates: from 01/02/2014 to 28/02/2014

Location: European Forest Institute, Joensuu, Finland

Host: Robert Mavsar, robert.mavsar@efi.int

Purpose of the STSM

The purpose of the STSM was to gain knowledge about international trade network model applied to NWFPs, through knowledge exchange with the host institution and other early stage researchers. The specific objectives were:

1. increase the knowledge on some key NWFPs, such as mushrooms and truffles (working group 1);
2. to study the state of the art of ComTrade databases (task force 2), and to clean the databases on the main NWFPs;
3. to analyze the international trade flows among countries and among geographical and economical groups of countries (task force 4);
4. to use Social Network Analysis (SNA) techniques to describe the structure of the international trade network, and to integrate some network indices in the “classical” time series analysis;
5. to write a paper draft which could highly contribute on the study of NWFPs trade flows.

This STSM meets the scientific aims of the COST Action FP1203 (European Non-Wood Forest Products Network) and contributes to the pan-European project “StarTree” on NWFPs, where I am involved for the development of the working package on the economical analysis of market flows.

Description of the work carried out during the STSM

During the four weeks of my STSM the work has been carried out at the European Forest Institute headquarters, in Joensuu (Finland).

The **first week** was totally dedicated to the selection of NWFPs for study and ComTrade databases cleaning. NWFPs have been classified by typology, and for each type we identified the respective HS codes. In the first selection we decided to study only wild products, and 75 databases have been downloaded by the ComTrade website (see table 1). But a further selection of the most significant NWFPs was necessary to start the initial investigation: so we decided to concentrate the analysis (and the cleaning procedure) on a few products:

- honey, an international widespread product;
- wild mushrooms, as the variety of codes allows the study of the different steps of the product processing in the supply chain;
- truffles, whose market has changed in the last years, and trade flows were not yet been subject to a systemic review;
- cork, as the structure of its trade network is centered on few market leading countries;
- edible nuts, with a focus on chestnuts, as they are produced in many European countries;
- tannins, whose databases are very small and easy to be used for cleaning tests.

Tab.1. First selection of NWFP, and their ComTrade HS codes.

NWFP types	ComTrade HS Codes	NWFP types	ComTrade HS Codes
Tar	380300; 380700	Essential oils	330129; 330190
Tanning	320110; 320120; 320130; 320190; 330210; 330290;	Exudate	330130; 380510; 380520; 380590; 380610; 380690
Phytochemical	293990; 293999	Foliage	060491; 060499
Sap	130219	Natural gums	130120; 130120; 130190; 380630
Bark products	450110; 450190; 450200; 450310; 450390; 450410; 450490	Honey and bee prod.	040900; 152190
Edible nuts	080212; 080221; 080222; 080231; 080232; 080240; 080250; 080290	Wild mushrooms	070951; 070959; 071159; 071232; 071233; 071239; 200310; 200390
Truffle	070959; 200310	Mosses & lichens	060410
Berries	081010; 081020; 081030; 081040; 081090; 081110; 081120; 081190; 081340; 200880; 200899	Medicinal and aromatic plants	090950; 091020; 091040; 091099; 121190; 130219
Live trees/plants	060210; 060220; 060491; 060499	Litter	060291

The **second week** was focused on SNA theory and on specific software's practice.

Collaborative learning was the adopted approach: each member of the research group contributed to the team knowledge through specific seminars, both theoretical and practical. We shared knowledge about:

- SNA theory, from network history to dynamic models;
- indicators for local scale governance evaluation, using social network analysis indices;
- nodeXL usage;
- UCINET usage.

The **third week** (and part of the fourth) was used for ComTrade databases cleaning.

The main steps are summarized in the following:

1. Creation of new columns denoting continent, sub-continent and membership of EU-28 (according to UN definitions).
2. Transformation of records to a clear "from-to" link, both with ISO-codes and with UN continents' subdivision, useful for SNA analysis.
3. World-code removal, as all the flows are reported doubly.
4. Mirroring the data and improving the veracity of the data.

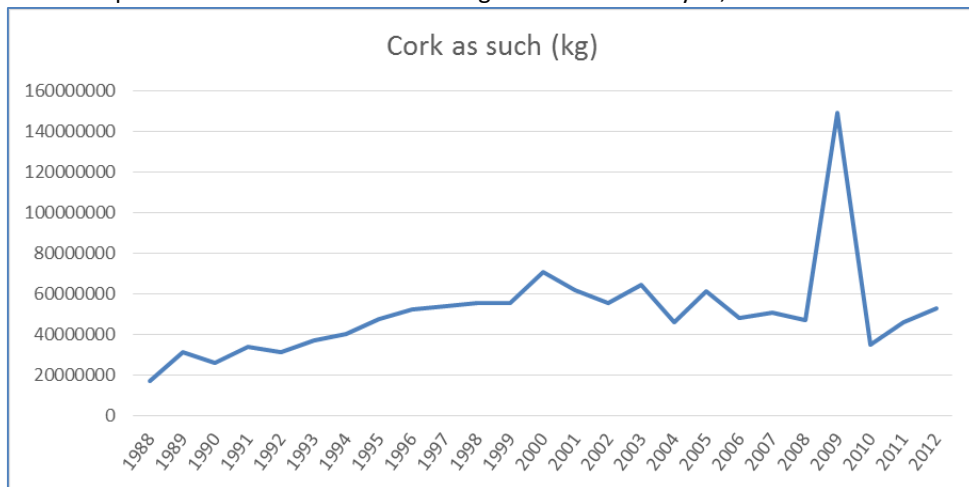
When the quantity is reported only by one country (most frequent option), the square matrix is filled with this value (mirroring the data).

In some cases, both the partners report the flow, but the two values are not symmetrical. When both the countries report data, but the quantities differ, we decided to use the highest quantities. The rationale is that NWFP are often underestimated, and we demonstrated that this approach minimizes the outliers for the selected NWFPs.

5. Double reporting check.
Flows were ordered by the main variables (commodity, year, partner, import/export flow), and duplicates were then deleted. Re-imports and re-exports have also been considered in this removal procedure.
6. Correction of outliers.
The general rule followed during the data cleaning was to keep values and quantities untouched, even if they appear quite unreliable. But in some cases the outliers were so evident that a correction was necessary. In the example showed in graph 1, the global trade of cork in 2009 looks to be three time higher than the other values in the time series. In this case it was necessary to

identify the origin of the error, and to substitute an estimation based on the quantities exchanged by the same two countries in the previous and following year.

Graph.1. Outliers identification through time series analysis, for cork as such.

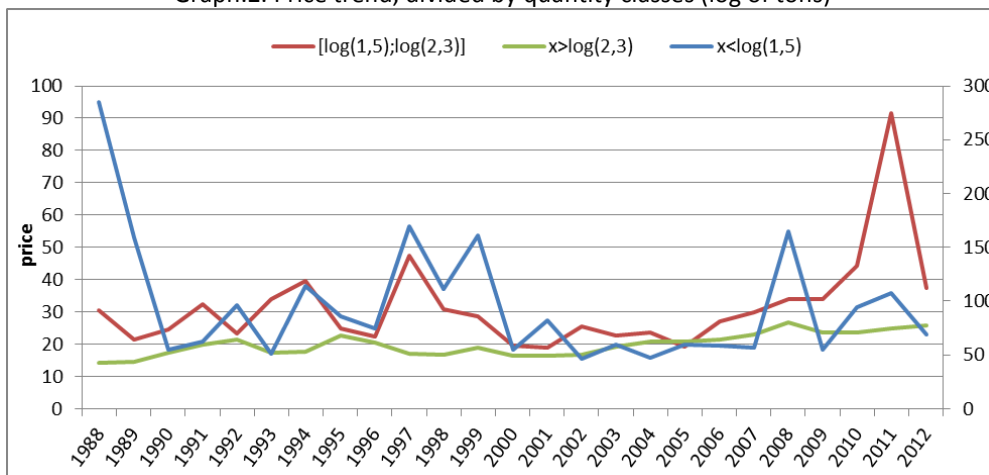


7. Estimation of missing quantities.

The majority of quantities are reported in kilograms. Exceptions are related to liters (already transformed in kg in the original dataset), and units of product: in all flows registered with this code, and in other missing data, the kilograms quantity is totally missing. Since the trade is reported with the value in USD, it has been decided to attempt an estimation of kilograms.

The percentage of missing quantities is quite low. For examples, in cork it is around 7,6%, equally distributed among the four categories belonging to this product. To estimate the missing kilograms, price per kg has been calculated in the rows where kilograms have been reported. We decided of not using the total average for the estimation, but on the other side we avoided to use too much stratification variables. After various attempts, flows have been divided by commodity and year, and then ordered by quantity. The price distribution showed that the ratio between USD and kg is more stable for higher exchanged quantity. For this reason, we divided data in three classes, and the thresholds have been decided looking at the stability of the trend. In the lowest two classes a weighted average has been adopted for estimation.

Graph.2. Price trend, divided by quantity classes (log of tons)



8. Selection of complete traceable flows for SNA. Further selection of data has been adopted for SNA: all the flows from or to an unknown country have been deleted.

The **fourth week** was committed to discussion of paper drafts and future work.

The initial idea was to prepare a paper draft focused on the description of the general market structure for key NWFPs, based on international trade flow analysis. At the end of the STSM period we were able to discuss in detail about two general paper drafts, and to outline other two more specific paper drafts.

Description of the main results obtained

The main aim of this STSM has been fulfilled: I was able to concentrate on international trade flows analysis applied to NWFPs, and to fill up my knowledge gaps on SNA theory and software's, thanks to the effective exchange with the other early stage researchers.

Besides the paper drafts, the main outputs of this month are two:

- an excel file with detailed instruction for ComTrade databases cleaning methodology;
- first outputs of the trade flows analysis for the main selected NWFPs, such as graphs and tables indicating the main importers and exporters of each commodity.

Future collaboration with host institution

The collaboration with the host institution is still ongoing.

We are now finishing the trade flows analysis for the main selected NWFPs, and the preparation of the final papers will allow a continuous debate in the future months.

Foreseen publications/articles resulting or to result from the STSM

The results of this STSM will be published in one or more peer-reviewed scientific journals. The articles that we are trying to submit to a journal in the next few months are at least three, two of which are already well defined.

The first paper will concern with the methodological work made during the STSM period. The draft title is: "Comparison of the data cleaning procedures for international trade data".

The idea is born by the fact that the UN ComTrade databases contain many problems related to data reliability, as (i) missing data on economic value or/and quantity and (ii) repetitive data entries (double counting), and there is not a common methodology to clean them. Data errors can thus lead to considerable errors in the estimation of the trade flows and analysis of market developments. To overcome this issue, we show how different approaches affect the results of typical trade analysis, and we try to propose a robust methodology for estimates. We'll finally discuss advantages and disadvantages of the different approaches, to allow the applicability with different commodities.

The second paper will focus on new challenges in economic networks, with a review of SNA applied to International Trade Network, and points to a clear demand on network explanations of economic concepts. The working title is: "Social Network Analysis as a tool for the analysis of international trade".

Traditional trade analysis mainly provide very basic indicators of trade, but they mostly fail to give a comprehensive picture of the market development and information on countries' position in the market, and relations between different countries. The paper will explain how network concepts/methods can deal with different problems in economic analysis of International Trade Network, and each problem will be addressed through an empirical example with different commodities.

Finally we are going to write specific papers explaining the market developments for different non-wood forest products, but we have not yet decided on which products we will specifically focus.