Geophysical Research Abstracts
Vol. 18, EGU2016-12510, 2016
EGU General Assembly 2016

Link prediction in the network of global virtual water trade

Marta Tuninetti, Stefania Tamea, Francesco Laio, and Luca Ridolfi
Politecnico di Torino, DIATI, Torino, Italy (marta.tuninetti@polito.it)

Through the international food-trade, water resources are 'virtually' transferred from the country of production to the country of consumption. The international food-trade, thus, implies a network of virtual water flows from exporting to importing countries (i.e. nodes). Given the dynamical behavior of the network, where food-trade relations (i.e. links) are created and dismissed every year, link prediction becomes a challenge.
In this study, we propose a novel methodology for link prediction in the virtual water network. The model aims at identifying the main factors (among 17 different variables) driving the creation of a food-trade relation between any two countries, along the period between 1986 and 2011. Furthermore, the model can be exploited to investigate the network configuration in the future, under different possible (climatic and demographic) scenarios.

The model grounds the existence of a link between any two nodes on the link weight (i.e. the virtual water flow): a link exists when the nodes exchange a minimum (fixed) volume of virtual water. Starting from a set of potential links between any two nodes, we fit the associated virtual water flows (both the real and the null ones) by means of multivariate linear regressions. Then, links with estimated flows higher than a minimum value (i.e. threshold) are considered active-links, while the others are non-active ones. The discrimination between active and non-active links through the threshold introduces an error (called link-prediction error) because some real links are lost (i.e. missed links) and some non-existing links (i.e. spurious links) are inevitably introduced in the network. The major drivers are those significantly minimizing the link-prediction error.
Once the structure of the unweighted virtual water network is known, we apply, again, linear regressions to assess the major factors driving the fluxes traded along (modelled) active-links.

Results indicate that, on the one hand, population and fertilizer use, together with link properties (such as the distance between nodes), are the major factors driving the links creation; on the other hand, population, distance, and gross domestic product are essential to model the flux entity. The results are promising since the model is able to correctly predict the 85% of the 16422 food-trade links (15% are missed), by spuriously adding to the real network only the 5% of non-existing links. The link-prediction error, evaluated as the sum of the percentage of missed and spurious links, is around 20% and it is constant over the study period. Only the 0.01% of the global virtual water flow is traded along missed links and an even lower flow is added by the spurious links (0.003%).