

Modeling the future evolution of the virtual water trade network: A combination of network and gravity models

Martina Sartori^{a,b}

Stefano Schiavo^{a,c,d}

Andrea Fracasso^{a,c}

Massimo Riccaboni^{e,f}

Abstract

The paper investigates how the topological features of the virtual water (VW) network and the size of the associated VW flows are likely to change over time, under different socio-economic and climate scenarios. We combine two alternative models of network formation – a stochastic and a fitness model, used to describe the structure of VW flows- with a gravity model of trade to predict the intensity of each bilateral flow. This combined approach is superior to existing methodologies in its ability to replicate the observed features of VW trade. The insights from the models are used to forecast future VW flows in 2020 and 2050, under different climatic scenarios, and compare them with future water availability. Results suggest that the current trend of VW exports is not sustainable for all countries. Moreover, our approach highlights that some VW importers might be exposed to “imported water stress” as they rely heavily on imports from countries whose water use is unsustainable.

Keywords: virtual water trade; complex networks; fitness model; agricultural production; preferential attachment; gravity model; water stress

JEL codes: F14, F18, Q25, Q56

Acknowledgements

Financial support received through the project “The global virtual-water network: social, economic, and environmental implications” (FIRB – RBFR12BA3Y) funded by the Italian Ministry of Education, University and Research (MIUR) is gratefully acknowledged. The authors are grateful to Francesco Laio and Stefania Tamea for providing the data on virtual water trade, and to Syud Amer Ahmed for sharing the projections on future GDP and population. Participants to the symposium “Acqua Risorsa Globale” (Torino 5-6 May 2016) provided useful feedback on an earlier draft of the paper.

^a University of Trento – School of International Studies.

^b IEFE – Bocconi University.

^c University of Trento – Department of Economics and Management.

^d OFCE-SciencesPo – DRIC

^e IMT – Lucca

^f KU Leuven – Department of Managerial Economics, Strategy and Innovation

1. Introduction

Virtual water (VW) is an indicator measuring the volume of water used to produce a good, implicitly “embedded” into a such good (Allan 1998) and represents a useful concept to identify the amount of water resources associated with international trade flows. Since its inception, the VW concept has inspired a flourishing literature on how to address global water scarcity vis-à-vis commodity production and consumption in a variety of disciplines. For a comprehensive review on this issue, see Antonelli and Sartori (2015).

The aim of this paper is to provide alternative projections about the future evolution of the international bilateral VW flows both in the medium- (2020) and in the long-run (2050), and to relate them to future water availability in order to assess their sustainability.

Several recent works have applied network analysis to investigate the properties and evolution of the network resulting from bilateral VW flows. We refer to two broad classes of network models to describe the network and to forecast its future evolution. The first is based on a preferential attachment mechanism (Barabási and Albert, 1999; Riccaboni and Schiavo, 2010) whereby nodes accumulate new links in proportion to the number of links they already have (a “rich get richer” mechanism). The second is a fitness-dependent model in which the ability to connect depends on an intrinsic characteristic of each node (a “good get richer” mechanism; see Garlaschelli and Loffredo, 2004; Suweis et al., 2011; Dalin et al., 2012).

We start by testing which models better reproduce the relevant topological characteristics of the VW trade network: this is done comparing the simulated network stemming from the models with the data on the basis of a number of standard topological measures. We then investigate the ability of the models to match the characteristics of the individual nodes. Adding this layer of analysis to the investigation on the global properties of the network represents, *per se*, a contribution to the literature.

Subsequently, we focus on the intensity of bilateral flows (as opposed to their mere existence), described by a gravity model of trade. This is an empirical framework whose ability to predict trade intensity is well-known in the economic literature (Anderson and van Wincoop, 2003; Head and Mayer, 2015).

Finally, we combine these models to produce projections about the future evolution of the network of VW flows in 2020 and 2050 under alternative scenarios. These projections are related to future water availability to assess the degree of water stress associated with the VW flows implied by the models. We find that the current trend of VW exports is not sustainable for all countries. Moreover, our network approach allows us to capture possible indirect effects: some importing countries may be prone to “imported water stress” as they rely on VW flows coming from countries whose trade patterns put high pressure on available water resources.

Our study improves and expands upon recent studies that have applied complex network analysis to study the topology (Sartori and Schiavo, 2015), the geography (Konar et al., 2011) and the temporal evolution of VW trade as a global network (Suweis et al., 2011; Carr et al., 2012; Dalin et al., 2012). First of all, our analysis is based a total of 309 agricultural goods up to the year 2010; second, new country characteristics are tested in the fitness model. Third, two alternative models are used to simulate the topology of the VW network and are compared in terms of their performance. Finally, the size of future VW flows is predicted by means of a gravity model of trade.

The paper proceeds as follows. The next Section briefly describes the data and the methodology used to replicate the binary structure of the VW network and the intensity of each flow. Section 3 describes the performance of the different models, whereas Section 4 builds on the results in order to generate future projections of VW flows in 2020 and 2050.

The last Section discusses the implications of these projections in terms of direct and indirect water stress.

2. The network of virtual water trade flows: data and methodology

The network representation of the global trade system is a graph made by N nodes (countries), connected by links that represent bilateral (virtual water) flows. The network is represented by a square matrix W_N (dimensions $N \times N$), where exporters are in rows and importers in columns. International trade gives rise to a weighted and directed network, in which the link direction goes from the exporting to the importing country, and the weight of each link is given by the volume of virtual water flowing between any country-pair. Each cell w_{ij} captures the VW flow from country i to country j , with $w_{ii} = 0$ by construction. The (horizontal) sum over row i is the total amount of VW exports of country i , while the (vertical) sum over column j is the total amount of VW imports of country j . We can derive a binary, unweighted version of the network by disregarding the information on link weights (i.e. the size of the flow) and simply accounting for the presence/absence of a trade connection. In this case, the $N \times N$ binary matrix A_N is called an adjacency matrix and its generic element a_{ij} is either one or zero depending on whether countries i and j are connected or not.

The VW content of a good is the volume of water that is used to produce it (Allan 1998). When a good is traded, its VW content is implicitly traded as well: VW trade thus refers to the cross-border flows of VW implied by international trade. A VW flow is obtained by multiplying the estimated country-specific VW content of each (agricultural) good by the registered volume of trade. Food production and international trade data for a total of 309 agricultural goods (i.e. crops and animal products) over the period 1995–2010 are obtained from the FAOSTAT database, while Mekonnen and Hoekstra (2011) provide estimates of the country-specific VW content of the various goods. Among existing studies, Konar et al. (2011) and Dalin et al. (2012) use a different method to determine the virtual water content of goods, namely the H08 global hydrological model by Hanasaki et al. (2008). For each single year, aggregate VW trade is obtained by summing the flows relative to the 309 individual goods. Our analysis includes 190 countries, listed in Table A1 of the Appendix.¹

Beside trade data, we also use information on the endowments of agricultural land (from FAOSTAT), long-term (average) annual precipitations, total renewable water resources (from AQUASTAT), GDP (in constant 2005 USD) and population (taken from the World Bank database). Projections about GDP and population are taken from the Shared Socioeconomic Pathways (Riahi et al., 2017).²

3. Replicating the observed network of virtual water trade

In this Section we employ a battery of models to replicate the main features of VW trade observed in the data. First, we use two simple models of network formation to describe the binary structure of VW trade, i.e. the presence or absence of a link between any two countries (Sect. 3.1 and 3.2). Second, we compare the relative performance of gravity and fitness models to describe the intensity of bilateral trade flows (Sect. 3.3). Building on these results, Section 4 will present projections about the future evolution of VW trade under the baseline scenarios

¹ For a detailed and exhaustive description of the way the virtual water content of the goods and the virtual water trade flows were computed, we refer to Carr et al. (2012, 2013) and Tamea et al. (2014). The Supplementary Material reports the list of products considered in computing the virtual water flows.

² The dataset is available at <https://tntcat.iiasa.ac.at/SspDb>. We use projections based on the OECD-SSP2 scenario labelled "middle of the road", that corresponds exactly to the medium variant of the new IIASA-VID-Oxford projections on population, which combines medium fertility with medium mortality and medium migration, and the Global Education Trend education scenario (see KC and Lutz, 2014, Dellink et al., 2015, and Riahi et al., 2017 for more details).

regarding the expected evolution of economic, demographic and environmental variables, whereas Section 5 will discuss the implications for this and other alternative scenarios illustrated in the Supplementary Material.

3.1 A stochastic model of network growth based on preferential attachment

Riccaboni and Schiavo (2010) develop a generalized version of the Barabási–Albert model to describe the dynamics and growth of weighted networks, and show that the model correctly replicates several features of international trade data. Although the VW trade network displays structural characteristics that are different from those analyzed in Riccaboni and Schiavo (2010), their modelling strategy is flexible enough to accommodate such differences. Since we focus on the binary network structure, the model boils down to a preferential attachment mechanism adjusted to allow for both entry of new nodes and the random allocation of some links.

The model starts with a given number of nodes (N_0) and assumes that at each time $t=\{1,\dots,M\}$ a new link among two nodes arises. Each new link is assigned to either existing or new nodes according to the following procedure:

- with probability a the new link is assigned to a new (i.e. not previously existing) source node, whereas with probability $(1-a)$ it is allocated to an existing node i ;
- in the latter case, the probability of choosing any existing node i is a linear combination of a random assignment and a preferential attachment mechanism. In particular, letting b identify the probability of assigning the new link randomly across existing nodes, the probability of choosing node i can be written as

$$p_i(t) = (1 - b) \frac{K_i(t-1)}{2t} + \frac{b}{N(t-1)} \quad (1)$$

where $N(t-1)$ is the number of active nodes at time $t-1$, and $K_i(t)$ represents the number of connection of each node (node degree).

The same process governs both sides of the link, meaning that source and destination nodes are treated symmetrically. By opportunely tuning the parameters a and b , as well as the number of initial nodes N_0 , we can obtain networks with very diversified topological structures (see the Supplementary Material for details).

The total number of links to assign equals the number of bilateral flows observed in the data, for each year in the period 1995-2010. To select the values of the model parameters (a and b) we run a grid search aiming at minimizing the Kolmogorov-Smirnov (K-S) test statistic that compares the degree distribution of the empirical and simulated networks. The latter is computed as the average across 1,000 replications of the (stochastic) procedure described above. The values of the K-S test referring to the year 2010 in the data are presented in Table 1. Given the stability of the network (see Sartori and Schiavo, 2015), results do not change substantially if we change the reference year.

The stochastic model that better replicates the actual patterns of VW trade features a combination of high entry rate a , mimicked by the initial number of nodes N_0 ($N_0=25$, about 10% of the final number of nodes), and a mixed preferential attachment hypothesis, where a randomly component coexists with the preferential attachment hypothesis (the probability that a new link is assigned randomly lies between 0.5 and 0.6). By fine-tuning the value of the parameter b , we find that $b=0.53$ is the value that minimizes the K-S test.

Table 1. Results from the K-S goodness of fit test for different entry rates and probabilities of random assignment. The number of initial countries N_0 is associated with the value of the parameter α : the higher N_0 , the closer the parameter α to zero. Red values indicate that the null hypothesis (i.e. empirical and simulated networks originate from the same distribution) is accepted at 5% significance level.

KS test		Number of initial countries (N_0)							
		25	50	75	100	125	150	175	200
Probability of random assignment (b)	0	0.6852	0.586	0.4849	0.3902	0.3	0.2034	0.1224	0.1178
	0.1	0.36	0.2986	0.2245	0.1533	0.0953	0.1285	0.1846	0.2425
	0.2	0.2276	0.1659	0.1103	0.114	0.1692	0.2374	0.2911	0.3425
	0.3	0.1458	0.0925	0.1187	0.1579	0.2243	0.2902	0.3556	0.4023
	0.4	0.0872	0.0879	0.1436	0.1853	0.2582	0.3284	0.3864	0.4243
	0.5	0.0472	0.093	0.1434	0.1972	0.2693	0.3391	0.3906	0.4271
	0.6	0.0545	0.0981	0.1479	0.2106	0.2751	0.332	0.3645	0.3883
	0.7	0.0759	0.107	0.1596	0.2108	0.2594	0.2933	0.3252	0.3547
	0.8	0.0966	0.1304	0.1626	0.1964	0.2301	0.2855	0.3355	0.4168
	0.9	0.1215	0.1593	0.1907	0.2397	0.2883	0.3439	0.4299	0.5589
1	0.1551	0.1907	0.2458	0.3009	0.3547	0.4463	0.579	0.8037	

3.2 The hidden variable hypothesis: a fitness-dependent model

A second popular model of network formation is represented by the so-call fitness model (Garlaschelli and Loffredo, 2004). In this case, nodes are assumed to have attributes that determine their ability to connect to others players. This “fitness” variable is defined as:

$$f_i = x_i / \sum_{j=1}^N x_j \quad (2)$$

where f is the definition of the fitness, x is the fitness quantity determining the structure of the network, and N the number of nodes. The fitness model is:

$$p(f_i, f_j) = \frac{\sigma f_i f_j}{1 + \sigma f_i f_j} \quad (3)$$

where $p(f_i, f_j)$ is the connection probability of every pair of nodes (i, j) of the simulated network, with $i \neq j$, and σ is a positive parameter of the model. The form of the connection probability $p(f_i, f_j)$ completely specifies the topological characteristics of the network. The parameter σ is determined by the following constraint:

$$L = 1/2 \sum_i \sum_{i \neq j} p(f_i, f_j) \quad (4)$$

where L is the observed value of the total number of links in the network. Equation (4) guarantees that the value of the expected total links in the simulated VW trade network equals the actual number of links.³

³ In order to obtain the simulated binary network of virtual water trade, the elements of the matrix $p(x_i, x_j)$, which are the probabilities of connection between each node obtained by applying equation (4), have been sorted in descending order. The first L elements have been replaced with value 1, meaning the presence of a link, while the remaining elements has been set equal to 0. In this way, the probability matrix $p(x_i, x_j)$ is transformed into the simulated binary matrix, whose total number of links equals the total number of links of the observed network.

Recent contributions by Suweis et al. (2011) and Dalin et al. (2012) use GDP to replicate the binary undirected VW network. Here, we test the performance of other variables associated with a country’s endowments, namely population (Pop), water endowment (WE), precipitations per capita (PrecPC) and agricultural land endowment (LE).

The similarity of the probability distributions of both the real and simulated node degree can be either confirmed or rejected by the K-S test (Table 2). The hypothesis that the empirical and the simulated networks originate from the same distribution is always accepted at the 5% significance level when the fitness variables are GDP and precipitations per capita, while it is never accepted in the case of land endowment. When the fitness model uses water endowment and population as fitness variables, the null hypothesis that the simulated and actual degree distributions are equal is rejected for 3 and 10 years respectively (over a total of 16). For this reason, in what follows, we shall consider the simulated networks using either GDP or precipitations per capita as fitness variable.

Table 2: Results of the K-S test on different fitness variables. The table reports for each year the result of a K-S test on the equality between the simulated and empirical degree distribution.

Year	Fitness Variables				
	GDP	Pop	LE	WE	PrecPC
1995	Accepted	Rejected	Rejected	Rejected	Accepted
1996	Accepted	Rejected	Rejected	Rejected	Accepted
1997	Accepted	Rejected	Rejected	Accepted	Accepted
1998	Accepted	Rejected	Rejected	Accepted	Accepted
1999	Accepted	Rejected	Rejected	Rejected	Accepted
2000	Accepted	Rejected	Rejected	Accepted	Accepted
2001	Accepted	Rejected	Rejected	Accepted	Accepted
2002	Accepted	Rejected	Rejected	Accepted	Accepted
2003	Accepted	Accepted	Rejected	Accepted	Accepted
2004	Accepted	Accepted	Rejected	Accepted	Accepted
2005	Accepted	Accepted	Rejected	Accepted	Accepted
2006	Accepted	Accepted	Rejected	Accepted	Accepted
2007	Accepted	Rejected	Rejected	Accepted	Accepted
2008	Accepted	Rejected	Rejected	Accepted	Accepted
2009	Accepted	Accepted	Rejected	Accepted	Accepted
2010	Accepted	Accepted	Rejected	Accepted	Accepted

Figure 1 displays the distribution of node degree observed in the data (for 2010) alongside those resulting from the simulations of the different network models (i.e., stochastic, GDP-based fitness, precipitations per capita-based fitness). We report both the complementary cumulative distribution (left panel) and the kernel density (right panel) to grasp the ability of the models to capture the connectivity distribution both in the body and in the tails.

To better assess our ability to match the structure of international VW trade, we compare the empirical and simulated networks by means of a set of popular measures that allow us to go beyond mere connectivity. We look both at network-wide indexes, such as density and assortativity, and node-specific features such as node degree, centrality and clustering, for which we take the average value across all nodes (see the Appendix for details on the computation of the indexes).

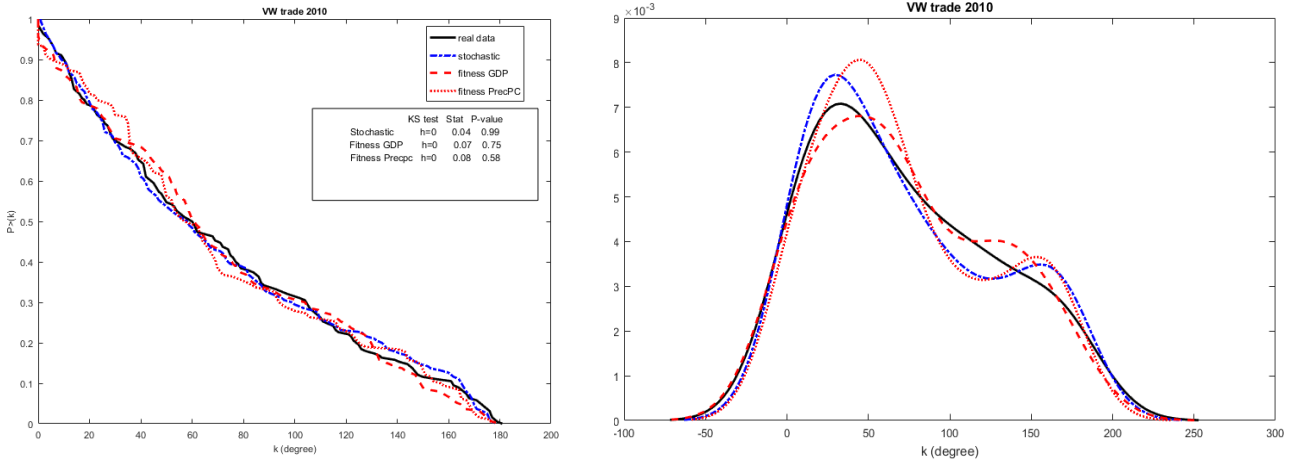


Figure 1. Left panel: Complementary cumulative probability distributions of empirical and simulated node degree obtained using both the stochastic and the fitness model. A textbox reports the results of the K-S test (statistic and P-value). Right panel: Kernel density of empirical and simulated node degree. Year 2010.

Table 3 collects and compares these indexes, computed for the empirical and the simulated networks obtained using both the stochastic and the fitness model.

The simulations for both the stochastic and the fitness model are designed in such a way that the number of nodes, links and density are mechanically replicated and therefore should not be used to evaluate the performance of the different models. A quick look at Table 3 suggests that the simulations do an excellent job in replicating average and maximum degree, as well as centralization. On the other hand, the actual values of assortativity are not very well replicated, whereas clustering lies somewhat in between. The stochastic model predicts lower levels of both assortativity (in absolute terms) and clustering, whereas the fitness model yields values that are higher (again in absolute terms) than what is observed in the data. Hence, an additional benefit of using both approaches is to derive some sort of upper and lower bounds to the actual behavior of the network that make results more robust. Last, we notice that the fitness models based on precipitations per capita and GDP have a very similar performance. These results hold over the entire period we consider (1995-2010), irrespectively of the specific year considered in the simulations.

Table 3. Empirical versus simulated networks: topology descriptive indexes and statistics. Year 2010.

VW Network	Nodes	Max degree	N° of links	Density	Assortativity	Centralization	Clustering
Empirical	190	181	13,719	38.2%	-0.27	0.58	0.71
Simulated: stochastic model	190	178	13,712	38.2%	-0.10	0.57	0.31
Simulated: fitness model (GDP)	190	181	13,719	38.2%	-0.56	0.58	0.82
Simulated: fitness model (PrecPC)	190	177	13,719	38.2%	-0.58	0.56	0.82

In our application, matching the aggregate structural properties of the empirical network may not be sufficient to evaluate a specific model. In fact, the overall performance of a model depends also on its ability to predict the attributes of individual nodes, such as node degree or

clustering.⁴ This kind of analysis has not been performed in previous studies using network models to generate predictions about the future development of virtual water trade (for instance Suweis et al., 2011; Dalin et al., 2012), and it therefore represents a further novelty of the paper.

Figure 2 displays a series of scatterplots that compare empirical and simulated values of out-degree obtained using either per capita precipitations or GDP as fitness variables for 1995 and 2010. This exercise turns out to be particularly informative: while the two models feature a similar performance in terms of aggregate network measures (see Table 3), the situation is quite different when we look at node-specific measures. In fact, the fitness model based on GDP generates a good match between the simulations and the data, the same does not apply to the model that uses per capita precipitations as fitness variable. In this case (left panels of Figure 2) there is no relationship at all between the out-degree predicted by the model and the values observed in the data.

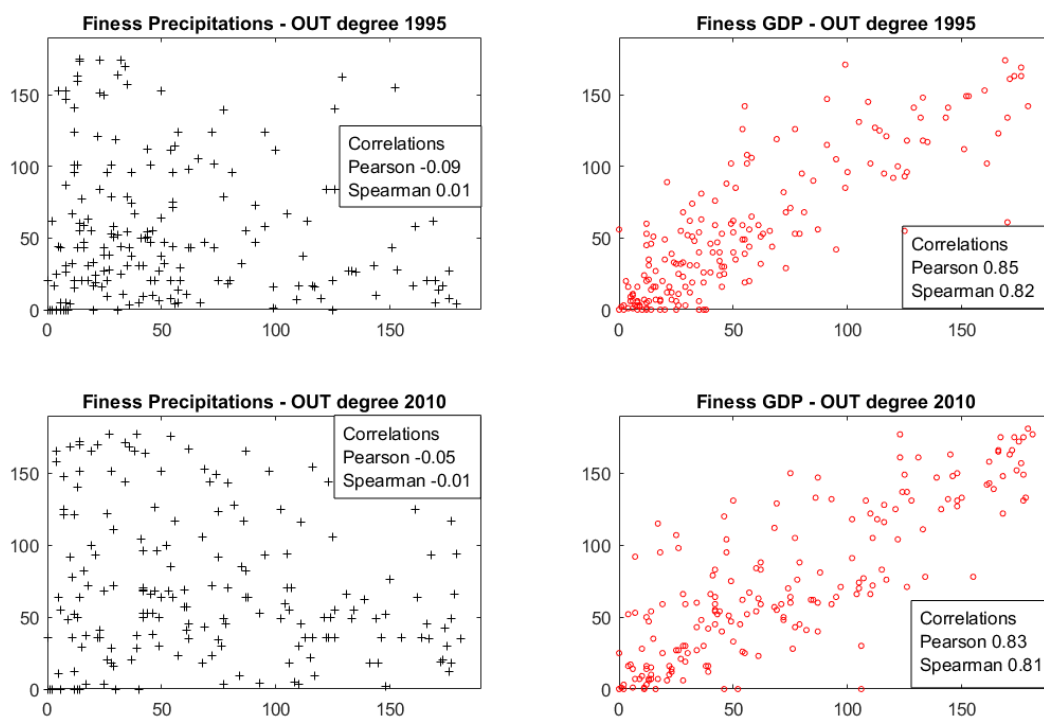


Figure 2. Scatterplot of simulated (vertical axis) versus empirical (horizontal axis) out-degree. Fitness model based on per capita precipitations (left panels) and GDP (right panels); data refer to 1995 (top) and 2010 (bottom). A textbox reports the values of Pearson’s and Spearman’s correlation coefficients.

Since the stochastic model does not identify nodes (they are all treated symmetrically), we match nodes to countries using connectivity, i.e. node out-degree. Hence, the node with highest out-degree is associated with the most connected country in the empirical network and so on for all nodes. As a result, there is a built-in correlation between empirical and simulated node out-degree and a scatterplot such as those in Figure 2 would not be particularly informative. However, a comparison between empirical and simulated values is still relevant for higher-order statistics such as clustering.

In Table 4 we report (Pearson’s and Spearman’s) correlation coefficients for out-degree and clustering for the three network models (fitness with precipitations per capita or with GDP,

⁴ We thank an anonymous referee for suggesting this type of analysis.

plus the stochastic model) in two different years. Results are stable over time, and the table displays values only for the first and last year in the sample. The table confirms that the fitness model based on GDP clearly outperforms the one using precipitations per capita as “hidden variable”, also in terms of clustering. On the other hand, the stochastic model displays relatively high levels of correlation for the clustering coefficient as well, indicating a good ability to replicate the behavior of individual nodes in addition to the good matching of aggregate measures described in Table 3.⁵

Table 4. Correlation coefficients among empirical and simulated node-specific network measures.

		out-degree			clustering coefficient		
		PrecPC	GDP	stochastic	PrecPC	GDP	stochastic
1995	Pearson	-0.09	0.85	0.99	-0.15	0.24	0.39
	Spearman	0.01	0.82	1.00	-0.20	0.35	0.49
2010	Pearson	-0.05	0.83	1.00	-0.22	0.32	0.19
	Spearman	-0.01	0.81	1.00	-0.31	0.47	0.54

This additional investigation leads us to exclude the fitness model based on precipitations per capita from the rest of the analysis: in fact, when using the models to discuss future scenarios of VW trade, it is crucial that they manage to identify precisely the behavior of individual countries, not only the aggregate structural properties of the network. This finding has also a more general implication as it challenges the practice, common in the literature, to assess network models by focusing exclusively on their ability to predict the aggregate topological properties of the network. When the behavior of individual nodes is relevant, this practice is potentially misleading, as node-specific features may vary a lot even within networks with similar topological features.

3.3 The gravity model: goodness of fit

Whereas in the Sections 3.1 and 3.2 we compare models of network formation able to describe the binary structure of VW trade, in this section we investigate the relative performance of a fitness model and of a gravity model in describing the intensity of the bilateral trade flows in the weighted network.

The gravity model is a workhorse empirical tool used in economics for describing international trade flows (see, e.g., Anderson and van Wincoop, 2003; Anderson, 2011; Head and Mayer, 2015). It relates bilateral flows of goods to the mass of the countries, their geographical distance and other (either pair- or country-specific) factors. Recently, this model has been applied to study international and regional flows of virtual water (Fracasso, 2014; Tamea et al., 2014; Fracasso et al., 2016).

The baseline, log-linear version of the gravity model can be written as:

$$\log X_{ij} = \alpha_0 + \alpha_1 \log GDP_i + \alpha_2 \log GDP_j + \alpha_3 \log Dist_{ij} + \varepsilon_{ij} \quad (5)$$

where the (log) exports from country i to country j (X_{ij}) are modelled as a log-linear function of the countries' GDP and their geographical distance ($Dist_{ij}$), plus a constant (α_0) and an error term (ε_{ij}). α_1 and α_2 measure, respectively, the elasticity of trade to the GDP of the source and destination country, whereas α_3 measures the elasticity of trade to distance. In this

⁵ Also in the case of clustering, to link simulated values obtained using the stochastic model with empirical values we first sort nodes in terms of out-degree, and then compute the correlation for the associated clustering coefficients.

study we adopt a slightly more elaborated version of this model by building on Fracasso (2014). In practice, to investigate the size of the bilateral virtual water flows (VW_{ij}), we estimate:

$$\log VW_{ij} = \beta_0 + \beta_1 \log GDPpc_i + \beta_2 \log GDPpc_j + \beta_3 \log Pop_i + \beta_4 \log Pop_j + \beta_5 \log Dist_{ij} + \beta_6 \log AraLand_i + \beta_7 \log AraLand_j + \beta_{ij} D_{ij} + u_{ij} \quad (6)$$

where GDP per capita ($GDPpc$) and population (Pop) capture economic mass, $AraLand$ measures the available endowment of arable land, $Dist$ is the geographical distance between the countries, and D_{ij} includes a set of dummy variables (taking value 1 when, respectively, countries have a common border, participate in a common preferential trade agreement, share the same official language, adopt the same currency, had the same colonizer post-1945). Data on the geographical distances between the most populated cities of each country-pair, as well as information about the dummy variables, are taken from the database maintained by the CEPII (see Head et al., 2010), while the sources of data on GDP, population and agricultural land are the same described in Section 2.

In order to deal with unobserved country-specific factors that might bias the estimates obtained by means of a cross-section analysis for one year, the model is estimated using panel data for the years 1995, 2000, 2005 and 2010, and it includes country-role specific fixed effects, in line with the recommendations of the most recent literature to reduce the risk of biased estimates for the coefficients of interest (Head and Mayer, 2015).

As anticipated, the predicting power of this gravity model applied to VW flows is compared against that of a fitness model calibrated to match the intensity of bilateral VW flows. In particular, we refer to the recent contribution by Dalin et al. (2012), who assume that VW trade responds to rainfall times agricultural area of the exporting countries, and to the population of importing ones. Labeling y_i and z_j the fitness of the exporting and importing country respectively, Dalin et al. (2012) assign each link a weight $w_{ij}(y_i, z_j) = \beta y_i z_j$, and calibrate β such that $2F = \sum_i \sum_{i \neq j} w_{ij}(y_i, z_j)$, F being the total VW flow observed in the system in a given period. This setup allows us to obtain a second set of predictions: comparing them to the ones resulting from the gravity model we can see which setup performs better.

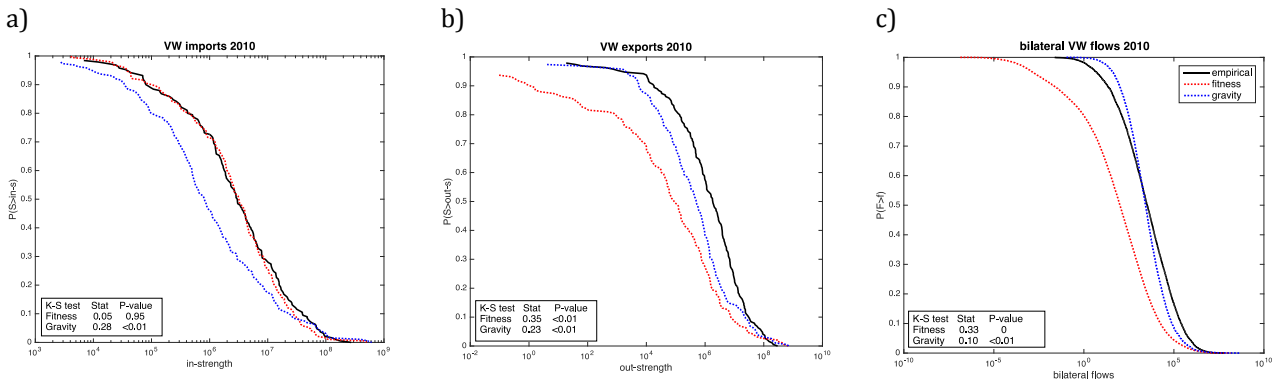


Figure 3. Complementary cumulative probability distribution of in-strength (panel a), out-strength (panel b) and bilateral VW flows (panel c). Comparison of empirical and simulated data at the year 2010; a textbox reports the results of the K-S test (statistic and P-value) for the two simulation models.

Figure 3 displays a comparison between the empirical and simulated distributions of countries' total imports (panel a), total exports (panel b), and bilateral flows (panel c) using the alternative approaches. A text-box reports the values of a K-S test for equality of the

gravity- and fitness-based distributions. Results confirm that the gravity model performs better in case of out-strength and bilateral trade flows (panel b and c). These results hold over the period we are considering (1995-2010), irrespective of the specific year considered. Hence, these findings suggest that the gravity model is a suitable approach to predict the volumes of VW flows and describe the so-called intensive margin of trade.

Once again, we are interested in looking beyond the aggregate behavior of the two models, and in assessing their ability to assign high-strength to countries that actually export large amounts of VW. Figure 4 presents scatterplots and correlation coefficients for actual and simulated values of out-strength (total export flows) for each country in the sample. Both models perform reasonably well and displays a positive correlation between simulated and actual values, but the gravity model features a much tighter fit, and higher values of correlation coefficients.

In what follows, therefore, we will use the gravity model to produce a set of out-of-sample projections on future VW flows, and combine them with either the fitness or the stochastic models of network formation, to determine the presence/absence of a link between any two countries. More precisely, we will use the estimated coefficients of the gravity model to deliver projections on the volumes of bilateral trade flows conditional on the existence of a link, as determined by either the stochastic or the fitness network model. These projections are based on the same prospective evolution for the underlying variables (i.e, a baseline scenario in Section 4 and alternative scenarios in the Supplementary Material).

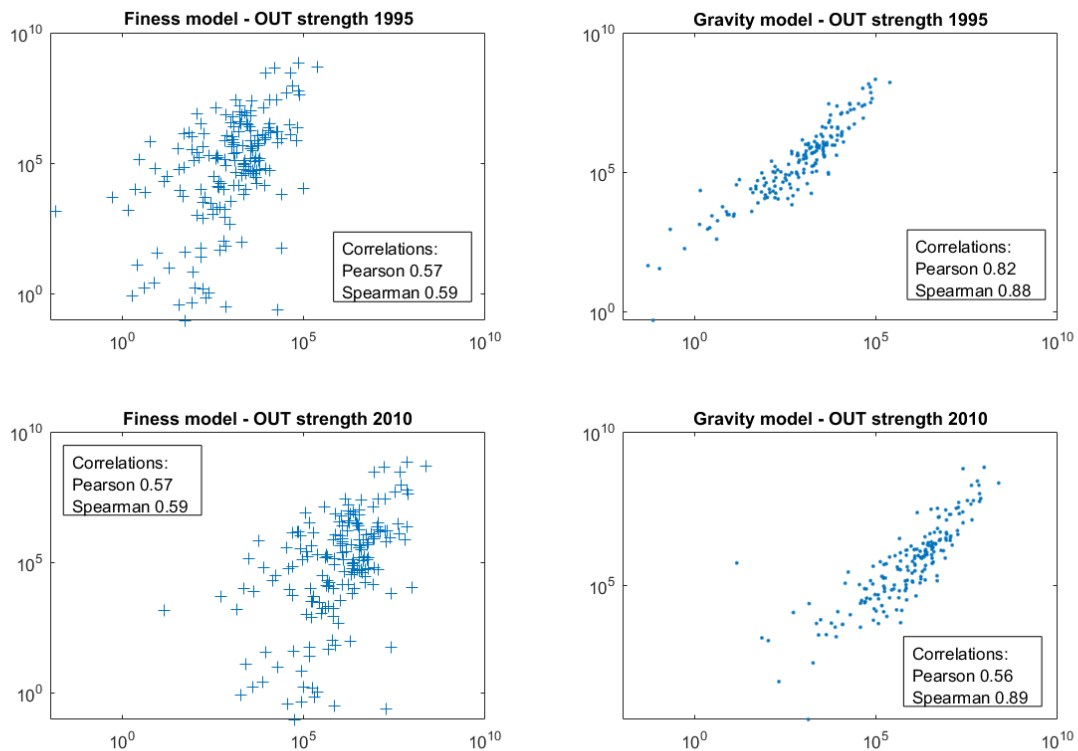


Figure 4. Scatterplot of simulated (vertical axis) versus empirical (horizontal axis) out-strength. Fitness model based on rainfall times agricultural area of the exporting countries, and the population of importing countries. Gravity model based on equation (6). Data refer to 1995 (top) and 2010 (bottom). A textbox reports the values of Pearson’s and Spearman’s correlation coefficients.

4. Future scenarios of virtual water trade

Taking stock of our previous results, we now present a set of projections about the future evolution of the VW network, along the lines of Suweis et al. (2011) and Dalin et al. (2012). We employ both a short-term scenario that uses predictions for 2020 and a long-term one calibrated on 2050. In this exercise, we are assuming that the structure of the system remains stable, so that the mechanisms that have generated the VW flows we observe in the data, will remain valid and drive the future evolution of the system. We are aware that this restriction trades off simplicity for realism, but this simplification is both necessary and common to all existing strategies used to forecast the future evolution of VW trade.

The topological structure of the network is predicted using two different network models (stochastic and fitness): hence, we end up with a total of four future scenarios (Table 5).

Table 5. The four scenarios for future VW flows

Model Family	Network model specification	Year	
		2020	2050
Stochastic	start from 2010 values	I	III
Fitness	GDP	II	IV

To predict the topological features of future scenarios of VW trade, we need information on the number of total links as well as on future projections of GDP, population and precipitations. To determine the number of future links present in the network (L^t), we follow Dalin et al. (2012), who estimate it on the basis of the average growth rate of the total links over the whole period considered in this analysis (1995-2010), which equals 1.95% in our context. The number of total links at the years 2020 (L^{2020}) and 2050 (L^{2050}) turn out to be 16,635 and 29,657 respectively. Future projections for GDP and population are taken from the SSP database described in Section 2 (Duval and de la Maisonnette, 2010; KC and Lutz, 2014; Dellink et al., 2015; Riahi et al., 2017); those for precipitations are available in the GAEZ database.⁶

4.1 Predicting the topological structure

To predict the evolution of the VW network with the stochastic model, we start from the empirical 2010 network and use the stochastic model to allocate the additional ($L^{2020}-L^{2010}$ or $L^{2050}-L^{2010}$) links. In this way we preserve the “identity” of each node and we minimize the risk to attribute a high degree to a small country or vice-versa. The key parameters governing the simulations (entry and randomness) are those reported in Section 3.1.

The complementary cumulative probability distribution functions (CCDF) of the logarithms of the simulated degrees at the year 2020 and 2050 are shown in Figure 5. The preferential attachment mechanism that drives the model implies that well-connected nodes capture a large share of new links, giving rise to a degree distribution with a thick upper tail and a minimum degree of about 90.

The fitness model provides us with another possible projection of the future structure of the VW trade network. Table 3 above suggests that it generates an upper bound for several of the topological features of interest that are instead underestimated by the stochastic model.

⁶ The GAEZ database (<http://www.fao.org/nr/gaez/about-data-portal/agro-climatic-resources/en/>) collects projections on several climate variables, including precipitation levels, obtained from a number of General Circulation Models (e.g., Hadley CM3, MPI ECHAM4, CSIRO Mk2, etc.) on the base of some SRES climate scenarios (A1F, A1, A2, B1 and B2), developed for the IPCC Fourth Assessment Report. We use projections on precipitations generated by the Hadley CM3 model, B2 SRES scenario, for the years 2020 and 2050 as our baseline, but also check the results using other climate scenarios, namely A1F1 and A2.

The procedure starts from assigning a value to the parameter of equation (3), σ , computed by fixing the future number of total links, $L^{2020}=16,635$ and $L^{2050}=29,657$. Figure 5 compares the CCDF of actual degree (years 1995 and 2010) and future scenarios at the year 2020 and 2050, using GDP as fitness variable. We observe that, if the total number of links grows at the average rate of the 1995-2010 time period (1.95% per year), the model predicts an increasing level of homogeneity in the distribution of node degree. In other words, the increasing density of the network implies that peripheral countries get more and more integrated into world trade, and thus catch up with top players in terms of number of trade partners.

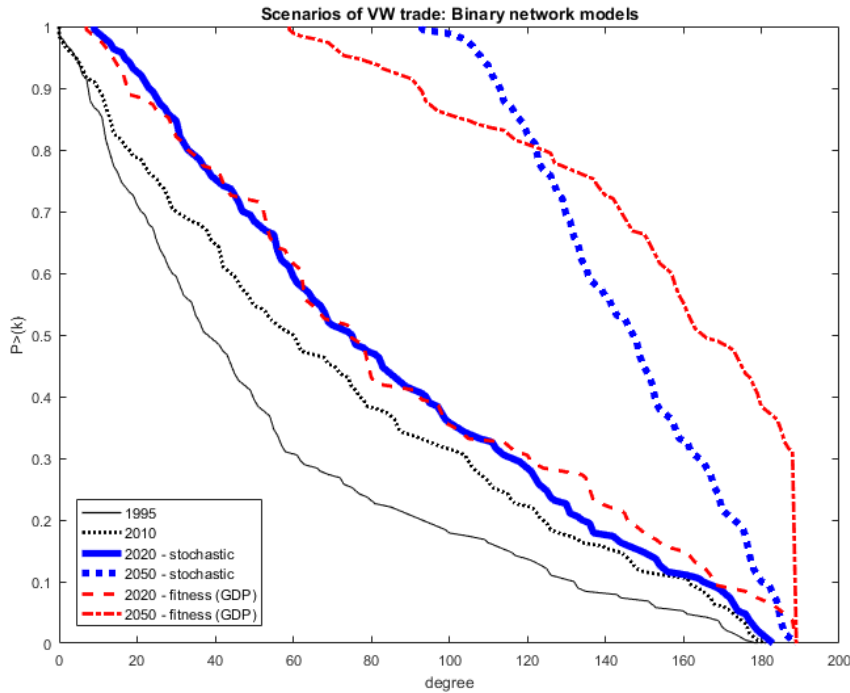


Figure 5. Complementary cumulative probability distribution function of the simulated out-degree for the year 2020 and 2050: stochastic model and fitness model. The figure also reports the empirical degree distribution for 1995 and 2010.

Consistently with other similar findings in the literature, Figure 5 suggests that poorly connected and peripheral countries will increase their relative importance in terms of VW trade. We see no significant differences across the projections for the short-term scenario of 2020, while the alternative models produce different results over a longer-term. These conclusions are confirmed by Kolmogorov-Smirnov tests that compare the cumulative distribution functions of the various models displayed in Figure 5 (see Table 6).

Table 6. Results for K-S test on the different models portrayed in Figure 5.

Year	Models compared	H	p-value	K-S stat
2020	Stochastic – fitness (GDP)	0	0.665	0.074
2050	Stochastic – fitness (GDP)	1	<0.01	0.311

The null hypothesis is the equality between the distributions of node degree obtained from the two models. H=1 indicates that the null hypothesis is rejected at 5% confidence level.

4.2 Predicting the intensity of VW flows

While the existing literature projects the evolution of both the binary and weighted structure of the VW network by means of the same (fitness) model (e.g. Suweis et al. 2011; Dalin et al. 2012), here we employ a gravity model to predict the intensity of each bilateral trade flow while taking the projected evolution of the VW network as calculated with either the stochastic model or the fitness model based on GDP.⁷

Using the estimated parameters from the regression of equation (6), thus, we predict the size of the virtual water flows in 2020 and 2050 by using the projections for GDP and population for those years, while keeping all the remaining variables constant. In particular, the endowments of arable land are treated as fixed, even though they might change due to urbanization, industrialization, demographic change, and technological change. Since different forces may have opposite effects on land availability, we take a neutral stance and prefer to consider it as fixed.

First, we predict bilateral VW flows for all possible country-pairs (meaning for all countries for which we have projections on GDP and population), then we only keep those links that are predicted by the network models.

Figure 6 reports the complementary cumulative probability distributions of node strength for the various scenarios under investigation: the left panel shows results obtained combining the stochastic and the gravity models, while the right panel is based on the fitness and gravity models.

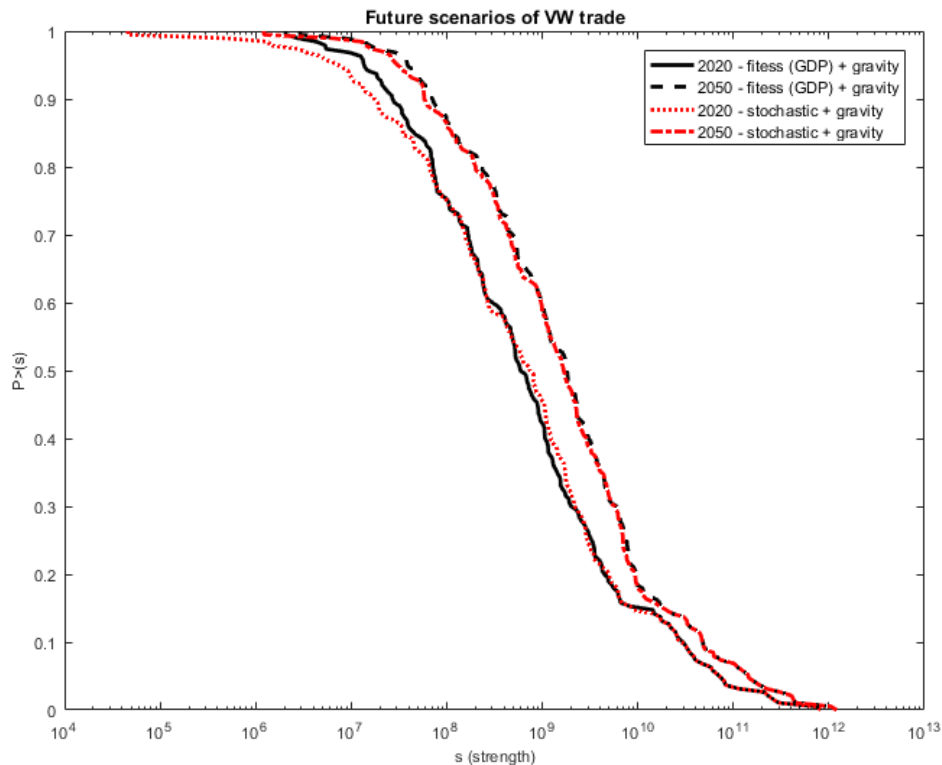


Figure 6. Complementary cumulative probability distributions of node out-strength (total VW export by country) in future scenarios (years 2020 and 2050). X-axis in log-scale.

⁷ Given the poor performance of the fitness model based on precipitations per capita to correctly predict node degree of individual countries, we only use the stochastic model and the fitness model based on GDP in this part of the analysis. It is worth noting that the fitness model based on precipitations per capita is the one that “shuts off” the largest number of bilateral links due to the simultaneous increase in population and reduction in precipitations in several locations as a result of climate change.

Overall, there are not big qualitative differences across the two families of models, which implies that these network models generate results that are broadly consistent with each other. Given the high uncertainty that surrounds forecasts about the relevant variables (especially long-term ones) we find the robustness of the results to the different network models rather reassuring. Indeed, in the case of the 2050 scenario, Kolmogorov-Smirnov tests (not reported) find not significant difference between the two distributions. Our projections suggest VW flows are bound to increase over the years, due both to the increased density of the network (as the number of links rises) and the stronger links between countries (for population and GDP are expected to grow in all countries). Over time we see that peripheral countries featuring small VW exports in 2010 increase their relevance in almost all scenarios. This is consistent with results in Sartori and Schiavo (2015), who find that the VW network has become more balanced over time, and with the predictions in Dalin et al. (2012).

5. Policy implications and concluding remarks

An important question stemming from the analysis is whether the projected VW flows are compatible with future water availability. Elaborating on the rainfall projections provided by the GAEZ database, we get an estimate of the future water availability by country. In short, we assume that future water availability will change in the same way as precipitations. This is tantamount to assume that changes in water availability are solely due to changes in precipitations, which is a very strong hypothesis. We are aware of the limits of this approach, especially when it comes to 2050: variations in water endowments depend also on other factors, such as the average annual temperature and the evapotranspiration (Gardner 2009), or the possibility that climate change brings about an acceleration of the water cycle.⁸ To (at least partly) account for the shortcoming of our approach, in what follows we discuss results based on different SRES scenarios. In particular, we use the B2 scenario as our baseline, but also consider scenarios A1F1 and A2 as an alternative. Moreover, for each scenario we consider both the average precipitation level and a confidence interval given by the average plus/minus one standard deviation.

Therefore, we apply the projected change in precipitations predicted by the GAEZ database to 2010 data on water availability by country. Then, we subtract from the estimated amount of future water availability the fraction of water resources that should be set aside for preserving environmental ecosystems, defined by the literature as Environmental Flow Requirement (EFR). A reasonable value for the EFR is about 30% of the total (blue) water availability (see for instance Smakhtin et al., 2004; Hirji and Davis, 2009). To account for all the remaining water-consuming activities of the economy, such as manufacturing, services and municipal water uses, an additional 30% of water resources is subtracted from the total water available.⁹

Each country's future net VW exports (i.e., exports minus imports of VW) computed from the network plus gravity models are then compared with the projected water availability. We classify each country according to the Water Stress Indicator (WSI), elaborated by Smakhtin et al. (2004). This indicator is computed as the share of the country's net VW exports over the total water availability, net of the fraction to set aside for preserving the environment and for non-agricultural water consumption. In this context, a value of the WSI larger than 1 indicates that the predicted amount of net VW exports is not sustainable as water resources would be

⁸ We thank an anonymous referee for pointing out this last factor.

⁹ It is estimated that, worldwide, agriculture accounts for about 70% of all water consumption, compared to 20% for industry and the 10% that is consumed for domestic use. At country level, however, differences in water allocation across sectors can be substantial (see Hoekstra and Mekonnen, 2012).

not enough to support the production of the associated agricultural goods; a value between 0.6 and 1 denotes that water resources will be heavily exploited; a value in the range 0.3—0.6 suggests that the projected net VW exports will moderately exploit water resources; finally, a WSI lower than 0.3 indicates that the production of agricultural goods for export will not generate a significant water stress. At this stage of the analysis it is important to point out that, despite the future available water is computed by setting aside the water needed for EFR and for non-agricultural water uses, the water stress indicator may possibly be underestimated for some agriculture-based economies, whose water use priority is for the agricultural sectors. Also, it is worth noting that the use of country-level variables may mask the presence of important regional disparities and situations of local water stress.

Results for 2050 (B2 SRES scenario, stochastic and gravity model) are illustrated in Figure 7, where grey areas refer to net VW importers while net VW exporters feature different colors based on the WSI (blue, light blue, orange and red in increasing order of water stress). Net VW exports are projected to exceed future water availability for agricultural purposes in some European countries, such as France, Denmark and the Netherlands. Other heavily stressed countries are likely to be some Eastern European (Ukraine) and Central Asian countries (Kazakhstan and Tajikistan).

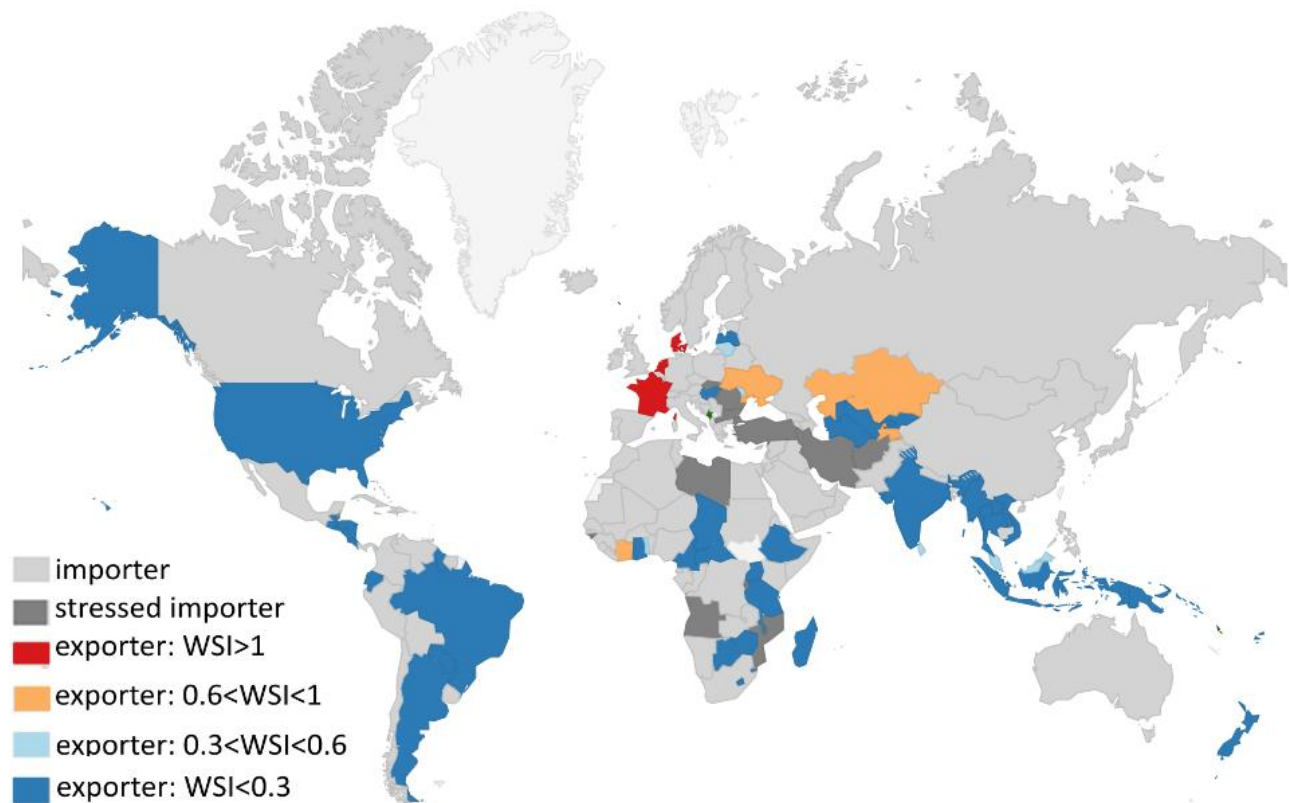


Figure 7. Predicted Water Stress Indicators (WSI) by country in 2050. Net VW exporters are blue if WSI is below 0.3; light-blue if WSI is between 0.3 and 0.6 (moderate exploitation of water resources); orange if WSI ranges between 0.6 and 1 (heavy exploitation of resources); red if WSI>1 (unsustainable situation). Dark grey regions represent net VW importers sourcing more than 50% of their imports from exporters with a WSI > 0.6 (stressed importers); light grey are other net VW importers. The map is generated using googleVis-0.6.2 and is based on results from a stochastic network model starting from the actual 2010 data, coupled with a gravity model to predict future bilateral VW flows. Future water availability is based on the Hadley global circulation model (Hadley CM3) under scenario B2. Maps obtained using different network models and different climatic scenarios are available in the Supplementary Material.

Different network models and different climatic scenarios deliver slightly different results (see the Supplementary Material for the complete set of maps), with the number of countries with an unsustainable position (“red countries”) ranging between 3 and 7.

Figure 7 further suggests that the majority of the water-scarce countries are net VW importers: this does not mean that they are not at risk. On the contrary, when these countries source a large fraction of their VW imports from countries whose export position is unsustainable, they may “import” water stress (for a recent contribution that stresses this point see Hoekstra and Mekonnen, 2016). We classify countries as “stressed importers” if they source more than 50% of their VW inflows from exporters with a predicted WSI > 0.6 . This is an admittedly crude way to capture the indirect effect of water stress in countries that may not be water-scarce per se, but rely extensively on VW imports. However, the importance of these indirect effects lies at the core of the network approach we have adopted in the paper.

Even this rough measure allows us to pinpoint a group of 16 importing countries that may come under “imported water risk” (in the terms of Hoekstra and Mekonnen, 2016) by 2050. Again, alternative scenarios change the number and identity of potentially stressed importers, whose number ranges between 11 and 17.

Within this group, some countries appear under most if not all scenarios: among these there are some small, import-dependent countries (such as Andorra and Malta), a number of African countries (Burundi, Rwanda, Guinea-Bissau and Mozambique), but also large countries such as Afghanistan, Romania and Turkey.

Different thresholds for identifying water stress and critical import shares would lead to the identification of diverse sets of stressed exporters and importers exposed to “imported water risk”: we leave the analysis of what thresholds are to be preferred as an issue for further research. Here, we limit ourselves to provide a *prima face* evidence showing that critical water conditions in certain countries are not just a matter of domestic interest but should be discussed at the multilateral level so as to take interdependences and spillovers into account.

The paper has investigated how the topological features of the virtual water network and the size of the associated VW flows are likely to change over time. The current trend of VW imports and exports, while broadly consistent with future water availability, is not sustainable for all countries. Technical change and trade agreements (two important drivers of water use and trade patterns) may help reducing the water requirements of several products, but demographic and dietary change may well put additional pressure on natural resources such as land and water. Incorporating these elements into the analysis represents a fruitful avenue for further interdisciplinary research that combines network analysis with insights from economics, agronomy, and climate science.

References

- Allan, J.A. (1998). "Virtual water: a strategic resource global solutions to regional deficits", *Groundwater*, 36: 545–546.
- Anderson, J.E. (2011). "The gravity model", *Annual Review of Economics*, 3: 133–160.
- Anderson, J.E. and van Wincoop, E., (2003). "Gravity with gravitas: a solution to the border puzzle", *American Economic Review*, 93: 170–192.
- Antonelli M. and Sartori M. (2015). "Unfolding the potential of the virtual water concept. What is still under debate?", *Environmental Science and Policy*, 50: 240-251.
- Barabási, A. and Albert, R. (1999). "Emergence of scaling in random networks", *Science*, 286: 509-512.
- Boguna, M. and Pastor-Satorras, R. (2003). "Class of correlated random networks with hidden variables", *Physical Review E*, 68: 036112.
- Caldarelli, G., Capocci, A., De Los Rios P. and Munoz, M.A. (2002). "Scale-Free Networks from Varying Vertex Intrinsic Fitness", *Phys. Rev. Lett.*, 89: 258702.
- Carr, J., D'Odorico, P., Laio, F. and Ridolfi, L. (2012). "On the temporal variability of the virtual water network", *Geophysical Research Letters*, 39: L06404, doi:10.1029/2012GL051247.
- Carr, J., D'Odorico, P., Laio, F. and Ridolfi, L. (2013). "Recent history and geography of virtual water trade", *PLOS ONE Public Library of Science*, 8: e55825. doi:10.1371/journal.pone.0055825.
- Dalin C., Suweis, S., Konar, M., Hanasaki, N. and Rodriguez-Iturbe, I. (2012). "Modeling past and future structure of the global virtual water trade network", *Geophysical Research Letters*, 39: L24402.
- Dellink, R., Chateau, J., Lanzi, E. and Magné, B. (2015). "Long-term economic growth projections in the Shared Socioeconomic Pathways", *Global Environmental Change*, doi:10.1016/j.gloenvcha.2015.06.004.
- Duval, R. and de la Maisonnette, C. (2010). "Long-run growth scenarios for the world economy", *Journal of Policy Modelling*, 32(1): 64-80.
- Fracasso, A. (2014). "A gravity model of virtual water trade", *Ecological Economics*, 108: 215–228.
- Fracasso, A., Sartori, M. and Schiavo, S. (2016). "Determinants of virtual water flows in the Mediterranean", *Science of the Total Environment*, 543: 1054-1062. <http://dx.doi.org/10.1016/j.scitotenv.2015.02.059>.
- Gardner, L.R. (2009). "Assessing the effect of climate change on mean annual runoff", *Journal of Hydrology*, 379: 351–359.
- Garlaschelli, D. and Loffredo, M.I. (2004). "Fitness-Dependent Topological Properties of the World Trade Web", *Physical Review Letters*, 93(18): 188701.
- Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K. (2008). "An integrated model for the assessment of global water resources – Part 2: Applications and assessments", *Hydrol. Earth Syst. Sci.*, 12: 1027-1037, doi:10.5194/hess-12-1027-2008.
- Head, K., Mayer, T. & Ries, J. (2010), "The erosion of colonial trade linkages after independence", *Journal of International Economics*, 81(1): 1-14.
- Head, K., and Mayer, T. (2015). "Gravity equations: workhorse, toolkit, and cookbook", in: Gopinath, G., Helpman, E., Rogoff, K. (Eds.), *Handbook of International Economics*, 4: 131–195. Elsevier.
- Hirji, R. and Davis, R. (2009). *Environmental flows in water resources policies, plans, and*

projects: Case studies, The World Bank Environment Department, Washington.

- Hoekstra A.Y., and Mekonnen M. (2012). "The Water Footprint of Humanity", *PNAS*, 109(9): 3232-3237.
- Hoekstra, A.Y. and Mekonnen, M. (2016). "Imported water risk: the case of the UK", *Environmental Research Letters*, 11(5): 055002.
- KC, S., and Lutz, W. (2014). "The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100", *Global Environmental Change*, <http://dx.doi.org/10.1016/j.gloenvcha.2014.06.004>
- Konar, M., Dalin, C., Suweis, S., Hanasaki, N., Rinaldo, A. and Rodriguez-Iturbe, I. (2011). "Water for food: The global virtual water trade network", *Water Resources Research*, 47, doi:10.1029/2010WR010307.
- Newman, M.E.J. (2001). "Clustering and preferential attachment in growing networks", *Physical Review E*, 64: 025102(R).
- Newman, M.E.J. (2003). "The structure and function of complex networks", *SIAM Review*, 45: 167-256.
- Riccaboni, M. and Schiavo, S. (2010). "Structure and growth of weighted networks", *New Journal of Physics*, 12(2): 023003.
- Riahi, K., van Vuuren, D., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Crespo Cuaresma, J., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L.A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A. and Tavoni, M. (2017), "The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview", *Global Environmental Change*, 42: 153-168, <http://dx.doi.org/10.1016/j.gloenvcha.2016.05.009>
- Sartori, M. and Schiavo, S. (2015). "Connected we stand: a network perspective on trade and global food security", *Food Policy*, 57: 114-127.
- Smakhtin, V., Revenga, C., and Döll, P. (2004). "A pilot global assessment of environmental water requirements and scarcity", *Water International*, 29(3): 307-317.
- Suweis, S., Konar, M., Dalin, C., Hanasaki, N., Rinaldo, A. and Rodriguez-Iturbe, I. (2011). "Structure and controls of the global virtual water trade network", *Geophys. Res. Lett.*, 38, L10403, doi:10.1029/2011GL046837.
- Tamea, S., Carr, J. A., Laio, F., and Ridolfi, L. (2014). "Drivers of the virtual water trade", *Water Resources Research*, 50: 17-28.

Appendix

Table A1. List of countries.

Afghanistan	Djibouti	Liberia	San Marino
Albania	Dominica	Libyan Arab Rep.	Sao Tome and Principe
Algeria	Dominican Rep.	Lithuania	Saudi Arabia
Andorra	Ecuador	Luxembourg	Senegal
Angola	Egypt	Madagascar	Serbia
Ant. and Bar.	El Salvador	Malawi	Seychelles
Argentina	Equatorial Guinea	Malaysia	Sierra Leone
Armenia	Eritrea	Maldives	Singapore
Australia	Estonia	Mali	Slovakia
Austria	Ethiopia	Malta	Slovenia
Azerbaijan	Faroe Islands	Marshall Islands	Solomon Islands
Bahamas	Fiji	Mauritania	Somalia
Bahrain	Finland	Mauritius	South Africa
Bangladesh	France	Mexico	Spain
Barbados	Gabon	Micronesia	Sri Lanka
Belarus	Gambia	Mongolia	Sudan
Belgium	Georgia	Montenegro	Suriname
Belize	Germany	Morocco	Swaziland
Benin	Ghana	Mozambique	Sweden
Bhutan	Greece	Myanmar	Switzerland
Bolivia	Grenada	Namibia	Syrian Arab Republic
Bosnia and Herz.	Guatemala	Nepal	Tajikistan
Botswana	Guinea	Netherlands	Thailand
Brazil	Guinea-Bissau	New Zealand	Macedonia
Brunei Darussalam	Guyana	Nicaragua	Timor-Leste
Bulgaria	Haiti	Niger	Togo
Burkina Faso	Honduras	Nigeria	Tonga
Burundi	Hungary	Norway	Trinidad and Tobago
Cambodia	Iceland	Oman	Tunisia
Cameroon	India	Pakistan	Turkey
Canada	Indonesia	Palau	Turkmenistan
Cape Verde	Iran (Islamic Rep.)	Panama	Tuvalu
Central African Rep.	Iraq	Papua New Guinea	Uganda
Chad	Ireland	Paraguay	Ukraine
Chile	Israel	Peru	United Arab Emirates
China, mainland	Italy	Philippines	United Kingdom
Colombia	Jamaica	Poland	United Rep. Tanzania
Comoros	Japan	Portugal	USA
Congo	Jordan	Qatar	Uruguay
Costa Rica	Kazakhstan	South Korea	Uzbekistan
Croatia	Kenya	Rep. of Moldova	Vanuatu
Cuba	Kiribati	Romania	Venezuela
Cyprus	Kuwait	Russian Federation	Viet Nam
Czech Republic	Kyrgyzstan	Rwanda	Yemen
Côte d'Ivoire	Lao Dem. Rep.	Saint Kitts and Nevis	Zambia
North Korea	Latvia	Saint Lucia	Zimbabwe
Dem. Rep. Congo	Lebanon	Saint Vinc. and the Gren.	
Denmark	Lesotho	Samoa	

Network Measures

Node degree $k_i = \sum_j a_{ij}$ is measures the number of contacts maintained by each node i , where a_{ij} is the element of the binary undirected adjacency matrix A_N . Density is the number of active links over their total possible number of links (which would occur if all nodes were connected with every other node).

Node strength, s_i , is the sum of all link weights w_{ij} of a node, and in the case of the VW trade network it represents the total VW imports (in-strength $s_i^{in} = \sum_j w_{ji}$) or total VW exports (out-strength $s_i^{out} = \sum_j w_{ij}$).

Assortativity is the tendency of highly connected nodes to be linked to other high-degree nodes. The assortativity index r for network G , computed following Newman (2001, 2003), is $r = \sum_{i,j \in G} (k_i - \bar{k})(k_j - \bar{k}) / \sum_{i \in G} (k_i - \bar{k})^2$, where k_i is the degree of node i and \bar{k} is the average degree in the network.

Centrality is meant to capture the position of each node within the network and its relative importance. Many centrality measures exist, here we compute the Freeman's network centrality: $fnc = \sum_i [D(n^*) - D(i)] / [(N - 1)(N - 2)]$, where $D(n^*)$ is the maximum degree of the degree distribution while $D(i)$ the degree of each single node, and N the number of nodes.

The clustering coefficient c_i represents the tendency of nodes to form tightly connected groups. Formally, $c_i = 2e_i / k_i(k_i - 1)$, where e_i is the number of links between the neighbors of node i and $k_i(k_i - 1) / 2$ is the maximum possible number of links existing between the k_i neighbors of i (Boguna and Pastor-Satorras, 2003). The clustering coefficient for the whole network is then given by the average value of c_i across all nodes.