Modeling past and future structure of the global virtual water trade network

C. Dalin,[1] S. Suweis,[2] M. Konar,[1] N. Hanasaki,[3] and I. Rodriguez-Iturbe[1]

Received 17 September 2012; revised 9 November 2012; accepted 13 November 2012; published 28 December 2012.

[1] Climate change and socio-economic development place an increasing pressure on essential natural resources, such as arable land and freshwater. The international food trade can save water globally by redistributing commodities produced relatively more water-efficiently. We focus on the global virtual water trade network associated with international staple food trade from 1986–2008. This study aims to determine which variables control the network's structure and temporal evolution, and to estimate changes in the network under future scenarios. Our fitness model reproduces both the topological and weighted characteristics of the network for the whole period. Undirected and directed network properties are well reproduced in each year, assuming as sole controls each nation's GDP, mean annual rainfall, agricultural area and population. The future structure of the network is estimated using climate and socio-economic projections, showing that volumes of virtual water traded will become increasingly heterogeneous and the importance of dominant importing nations will further strengthen. **Citation:** Dalin, C., S. Suweis, M. Konar, N. Hanasaki, and I. Rodriguez-Iturbe (2012), Modeling past and future structure of the global virtual water trade network, *Geophys. Res. Lett.*, *39*, L24402, doi:10.1029/2012GL053871.

1. Introduction

[2] Water resources are unevenly distributed on the planet. In some regions, while population grows and diets shift toward water-intensive products like meat, water resources are placed under increased pressure, leading to water and food security issues. Besides, many areas of the world are expected to suffer increasingly frequent and intense droughts under climate change, which will strain water resource use in agriculture even more and potentially lead to crop failures [*Field et al.*, 2012]. However, other regions have abundant water resources, prosperous agriculture and might slightly benefit from climate change in terms of crop yields [*Intergovernmental Panel on Climate Change*, 2007]. Thus, among different strategies to increase agricultural water-use efficiency (i.e., mechanization, water-saving irrigation, fertilizers, etc.), the international trade of agricultural products is a way to improve global water-use efficiency by virtually transferring water resources to water-stressed regions.

[3] The international food trade has proven to save water [*de Fraiture et al.*, 2004; *Yang et al.*, 2006], and to increasingly do so over time [*Dalin et al.*, 2012; *Konar et al.*, 2012], as countries with low agricultural water-use efficiency tend to import food from more efficient countries. These global water savings exist even though water is usually not a strong driver of international trade. Thus, it is important to understand the major factors controlling water transfers through international food trade, or global virtual water trade.

[4] Complex network analysis has revealed key properties of the global virtual water trade network based on staple food (GVWTN), such as a wide heterogeneity of the network's flows [*Konar et al.*, 2011; *Suweis et al.*, 2011; *Carr et al.*, 2012], a power law relationship between each nation's number of trade partners and volume of virtual water traded [*Konar et al.*, 2011; *Suweis et al.*, 2011; *Dalin et al.*, 2012], the tendency of nations that trade large amounts of water to connect to each other [*Konar et al.*, 2011] and a community structure [*D'Odorico et al.*, 2012], among others. *Suweis et al.* [2011] built a fitness model that reproduces the global structure of the GVWTN in 2000 using a few controlling variables: national GDP, rainfall and agricultural area. This model can be used to predict the future global structure of the network under climate and socio-economic changes.

[5] In this Letter, we built the GVWTN associated with five major crops and three livestock products from 1986–2008. We then compare the structure of the empirical network each year with the structure obtained from the fitness model using input variables specific to each year. Finally, we show the predicted structure of the GVWTN under socio-economic and climate scenarios. This analysis aims to answer the following questions: Can our theoretical model capture network properties in the 1986–2008 period? How can this model reproduce the directed structural features of the GVWTN over this period? What are the consequences of projected climate and socio-economic changes on the structure of the GVWTN in the coming decades?

2. Building the GVWTN Over Time

[6] In the global virtual water trade network associated with staple food trade (GVWTN), each node represents a country and each link between a pair of nodes is directed by the direction of trade and weighted by the volume of virtual water involved in the traded commodities. The main network measures are the node degree, k, referring to the number of connections of a node, and the node strength, s, referring to the sum of the weights assigned to each node's links. Thus here, the node degree is the number of trading partners of

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[2]Department of Physics, University of Padova, Padua, Italy.
[3]National Institute for Environmental Studies, Tsukuba, Japan.

Corresponding author: C. Dalin, Department of Civil and Environmental Engineering, Princeton University, Engineering Quad, Princeton, NJ 08544, USA. (cdalin@princeton.edu)

0094-8276/12/2012GL053871

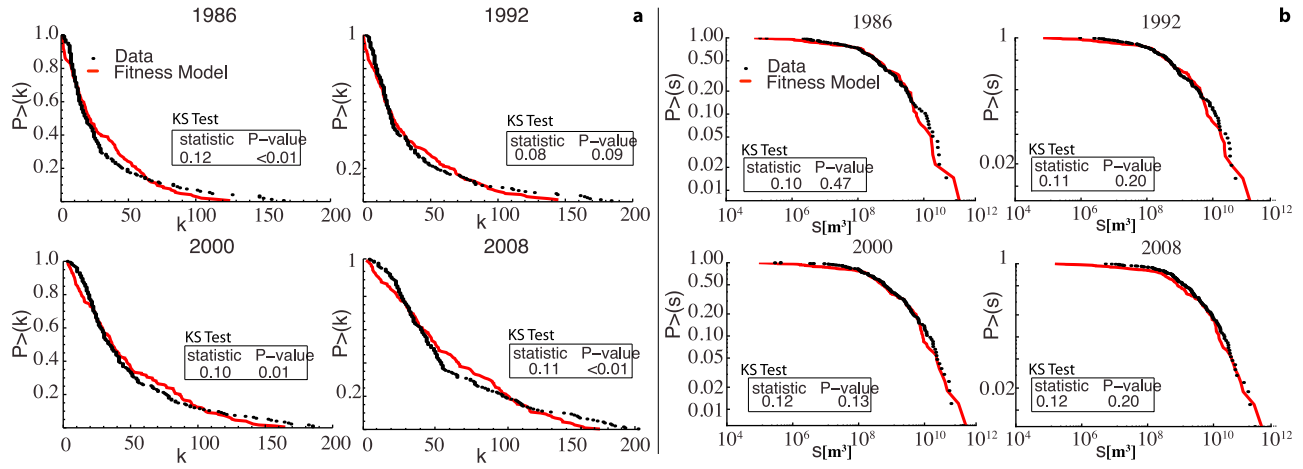

Figure 1. Exceedance probability distribution of the undirected node (a) degree (k) and (b) strength (s): comparison of data and model results in 1986, 1992, 2000 and 2008, using *GDP* and *RAA*, respectively. The similarity between data and model is confirmed by a Monte-Carlo Kolmogorov-Smirnov test; test results are shown in the "KS test" box.

each country and the node strength is the volume of virtual water traded by each country.

[7] We built the GVWTN associated with 58 food commodities made from five major crops (barley, corn, rice, soy and wheat) and three livestock products (beef, pork and poultry) from 1986–2008 using international food trade data [*Food and Agriculture Organization of the United Nations* (*FAO*), 2010a] and products Virtual Water Content (VWC) simulations from the H08 global hydrological model [*Hanasaki et al.*, 2008].

[8] Each link weight is obtained by multiplying the traded volume of a specific commodity by the VWC of this commodity in the exporting country: $VWT^n_{i,j,x} = VWC^n_{i,x} \cdot T^n_{i,j,x}$ where $VWT^n_{i,j,x}$ (kg_{water}) is the volume of virtual water exported from country i to country j through trade of commodity x in year n, $T^n_{i,j,x}$ ($kg_{product}$) is the volume of commodity x exported from i to j in year n and $VWC^n_{i,x}$ ($kg_{water}/kg_{product}$) is the virtual water content of commodity x produced in country i from all sources of water in year n (see auxiliary material).[1] In this study, we analyze the aggregated GVWTN, built by summing virtual water trade from selected commodities (representing about 60% of global calorie consumption [*FAO*, 2010b]) in a given year, in its directed and undirected versions.

3. Structure and Controls of the GVWTN Over Time

3.1. Undirected Network

[9] In this section, we focus on the undirected version of the GVWTN, in which a node's degree is the number of export and import partners of each country and a node's strength is the nation's volume of virtual water exports and imports. We find that the distribution of node degree is well fit by an exponential distribution every year. Similarly, the distribution of node strength is well fit by a stretched exponential distribution every year, presenting larger heterogeneity than the node degree distribution.

[10] To model the nodes' degree and strength, we attribute to each node two fitness variables: one based on the country's annual GDP (time series from 1986–2008 [*The World Bank*, 2012]) and the other on the nation's long-term average annual rainfall [*FAO Aquastat*, 2010] and agricultural land area (time series from 1986–2008 [*FAO*, 2010c]). The two fitness variables assigned to each node i, normalized *GDP* and normalized Rainfall times Agricultural Area (i.e., *RAA*), were used to characterize the undirected network properties in 2000 [*Suweis et al.*, 2011] and are defined as follows for years n = 1986–2008 and countries i = 1–N: $x^n_i = GDP^n_i / \sum_{j=1}^{N} GDP^n_j$ and $y^n_i = RAA^n_i / \sum_{j=1}^{N} RAA^n_j$. We then construct the network connections and flows with the following modeling process: for each year n, we connect every pair of nodes (i, j) (where $i \neq j$) with a probability $p\left(x^n_i, x^n_j\right) = \frac{\sigma^n x^n_i x^n_j}{1 + \sigma^n x^n_i x^n_j}$; and assign a weight $q\left(y^n_i, y^n_j\right) = \beta^n y^n_i y^n_j$ to each link between i and j, representing the virtual water trade flow between these nodes. The two parameters of the model, σ^n and β^n, are determined by the following consistency conditions: $1/2 \sum_i \sum_{j \neq i} p\left(x^n_i, x^n_j\right) = L^n$ and $1/2 \sum_i \sum_{j \neq qi} q\left(y^n_i, y^n_j\right) = F^n$, where L^n and F^n are the observed values in year n of the network's total number of links and total flux, respectively (see auxiliary material). As this theory relies on principles for the ensemble of nodes, this modeling framework applies to the complete network only. Both for node degree and node strength, we observe that the model and data match closely for all years studied (Figure 1). The similarity between the analytical and empirical results is confirmed by a Monte Carlo based Kolmogorov-Smirnov test [*Keutelian*, 1991; *Lilliefors*, 1967]: the hypothesis that the model and the data have the same distribution is accepted at the 1% significance level for k_{un} in most years and at the 5% level for s_{un} in all years (Figure 1; see auxiliary material).

[1] Auxiliary materials are available in the HTML. doi:10.1029/2012GL053871.

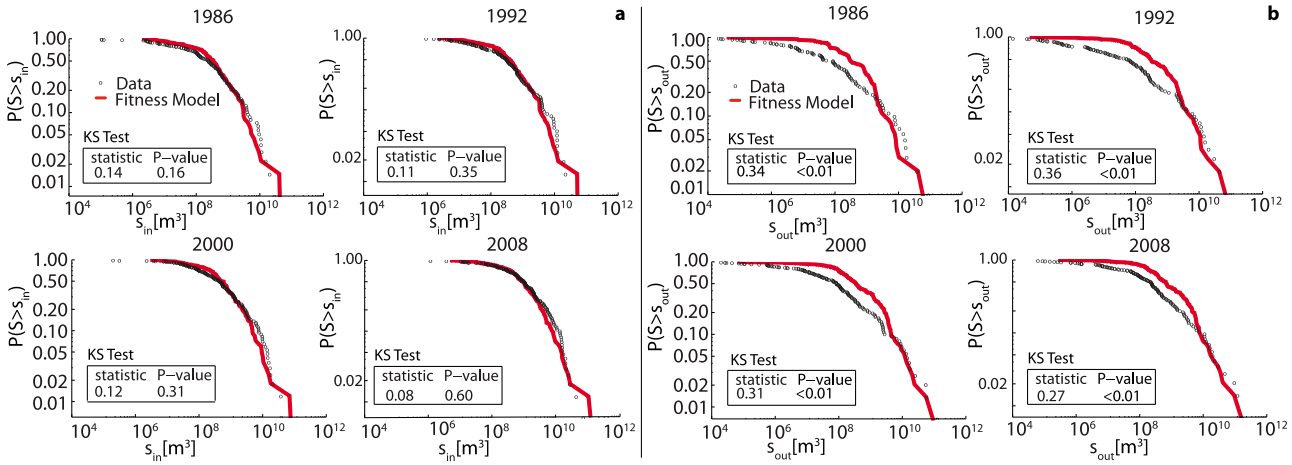

Figure 2. Exceedance probability distribution of the node (a) in-strength (s_{in}) and (b) out-strength (s_{out}): comparison of data and model results using *Population* and *RAA* in 1986, 1992, 2000 and 2008.The similarity between data and model is confirmed by a Monte-Carlo Kolmogorov-Smirnov test for s_{in} but not s_{out}; test results are shown in the "KS test" box.

3.2. Directed Network

[11] In this section, we analyze the directed version of the GVWTN, in which each node has two weighted properties: in-strength, s_{in} (i.e., import volume), and out-strength, s_{out} (i.e., export volume). We model these major directed properties for the first time. The fitness model framework reproducing the undirected network flows has been applied to the directed network [*Suweis et al.*, 2011], identifying unidirectional and bidirectional trade between each country pair, but without distinguishing between the strength sequences $\left(s_{in}^{i}\right)_{i=1..N}$ and $\left(s_{out}^{i}\right)_{i=1..N}$. However, directed flows are crucial to analyze the GVWTN. To characterize the asymmetry of the directed and weighted network, the fitness variable assigned to each node needs to be different whether the country participates in the trade as an exporter or as an importer. Indeed, even though each individual trade flow corresponds to the same volume for the importer and exporter ($s_{in}^{i,j} = s_{out}^{j,i}$), the directed strength sequences are different ($s_{in}^{i} \neq s_{out}^{j}$) since the node strength represents trade flows aggregated at the country level.

[12] Thus, in this paper, we choose for the first time two different fitness variables for exporting and importing nodes. As *RAA* represents well the nation's potential to produce and thus export food, we use this variable for exporting nodes, and for importing countries, we choose a variable related to food demand, represented by the national population. We assign to exporting nodes i a fitness variable using *RAA* as follows: $y_i^n = RAA_i^n / \sum_j RAA_j^n$ and to importing nodes i a fitness variable using population (*Pop*) as follows: $z_i^n = Pop_i^n / \sum_j Pop_j^n$ in each year n. We then attributed to each link directed from i to j a weight $\vec{q}\left(y_i^n, z_j^n\right) = \gamma^n y_i^n z_j^n$. The parameter γ^n is determined by the following consistency condition: $\sum_i \sum_{j \neq i} \vec{q}\left(y_i^n, z_j^n\right) = F^n$, where F^n is the observed total flux (see auxiliary material) in the network in year n (n = 1986 − 2008).

[13] Using this procedure, we obtain a close match between the modeled and observed distributions of directed strengths over the period (Figure 2), except for an overestimate of medium s_{out} values. The similarity between model and data is confirmed by the same test at the 5% level in all years for s_{in} but not for s_{out}. Thus, simple node properties such as rainfall, agricultural area and population allow us to model the two distinct distributions of exports and imports in the GVWTN, however the model accuracy remains to be improved for export flows.

4. Future Scenarios

4.1. Undirected Network

[14] Having shown the good performance of the fitness model to reproduce the statistical properties of the GVWTN using a few external variables, we apply this model to predict the future structure of the network under different scenarios for the year 2030. To predict the network's topological features, we used projections of national *GDP* for 2030 [*FAO*, 2002] and built the fitness function $p\left(x_i^T, x_j^T\right) = \frac{\sigma^T x_i^T x_j^T}{1+\sigma^T x_i^T x_j^T}$; where x^T is the projection of each nation's fitness variable (i.e., normalized *GDP*) in year T = 2030. The parameter σ^T is determined by estimating the future total number of links L^T. Using the time series of the mean degree $\langle k \rangle$ from 1986–2008, we estimate two future values of k^{2030} based on the average growth rate over the whole period (1986–2008) and over the 5 years 2003–2007: 3.6%/yr and 0.66%/yr, respectively. These rates lead to very different predictions, as the interconnectivity of the network grew significantly from 1986–2001 and much slower after 2002. In the slow growth case, the estimated mean degree in 2030 is about $\langle k \rangle^{2030}$ = 63 trade partners/nation, i.e., L^{2030} = 5, 777 links, while in the fast growth case, it is about $\langle k \rangle^{2030}$ = 118 trade partners/nation, i.e., L^{2030} = 10, 874 links (vs. 54 trade partners/nation and 5,402 links in 2008).

[15] We observe that, in the slow growth case, projections of *GDP* in 2030 do not result in a significantly different node degree distribution from that observed in 2008 (Figure 3a, red curve). However, assuming that the mean degree keeps growing at the average rate of the 1986–2008 period, the model predicts a much more homogeneous distribution of

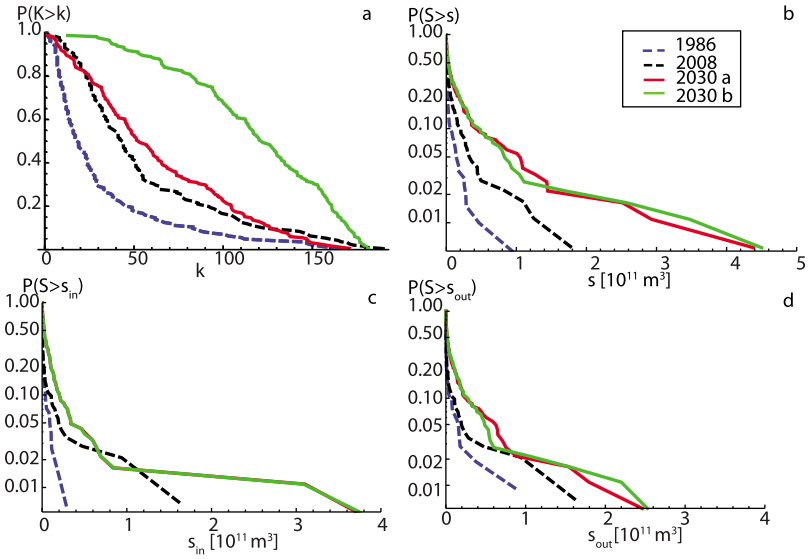

Figure 3. Future scenarios for undirected (a) degree (k) and (b) strength (s), (c) in-strength (s_{in}) and (d) out-strength (s_{out}). Distributions observed in 1986 and 2008 and predicted by the model for 2030 in different scenarios. Panel a: the number of links in the network is assumed to grow by 0.66%/yr (as from 2002–2007, red) or by 3.6%/yr (as from 1986–2008, green). Panel b-d: distributions predicted by the model with a constant growth of VWT volume (4.6%/yr as from 1986–2008) and two climate change scenarios: driest (red) and most humid (green).

node degree with a general increase of nodes' connectivity (Figure 3a, green curve). This result makes sense, as a rapid growth of the network's connectivity implies that poorly connected nodes will increase their degree and a node's degree cannot exceed the total number of nodes (about 184 nations), which limits the extent of the node degree distribution's tail and heterogeneity.

[16] To predict the network's weighted features, we used projections of rainfall (A2 scenario of the WCRP-CMIP3 for 2030–2050 [*Meehl et al.*, 2007]) and agricultural area [*FAO*, 2002] for 2030 and built the fitness function $q\left(y_i^T, y_j^T\right) = \beta^T y_i^T y_j^T$ where y^T is the projection of the fitness variable (i.e., normalized *RAA*) in year T = 2030. The parameter β^T is determined by estimating the future total VWT flow F^T. Using the time series of F from 1986–2008, we estimate a future value F^{2030} based on the average growth rate of the mean strength $\langle s \rangle$ over the whole period (1986–2008): 4.6% per year. The volume of VWT in the network has grown significantly from 1986–2008 and, unlike the interconnectivity, did not present any stabilization in the most recent years. Indeed, the total volume traded is limited in theory (i.e. by the global volume of food production) but it is unlimited in practice, as the limit is far from being reached (less than 1% of global cereal production was traded internationally in 2008 [*FAO*, 2010a]). We thus use the average rate over the whole period (1986–2008): 4.6%/yr, which corresponds to a mean strength of $\langle s \rangle^{2030}$ = 18.4 km^3, i.e., F^{2030} = 1, 700 km^3 (vs. $\langle s \rangle^{2008}$ = 6.7 km^3 and F^{2008} = 671 km^3).

[17] To run comparative scenarios, we selected the driest and most humid rainfall projections from the ensemble of climate projections WCRP-CMIP3 [*Meehl et al.*, 2007], which predicts national rainfall depth. We observe that the difference between projected node strength distributions with the driest and with the most humid climate scenario appears to be negligible. Both rainfall projections correspond to an increased strength of all countries, particularly for dominant nations (Figure 3b). Indeed, the projected strength distributions present remarkably large tails, implying that the dominance of a few countries in the GVWTN will increase and the heterogeneity among nations will grow.

[18] Global rainfall over land is predicted to decrease by 6% in the most humid model and 23% in the driest model from 2008–2030 [*Meehl et al.*, 2007], with large differences in trends among regions - rainfall will increase in some regions and decrease in others -, while global agricultural area is predicted to slightly increase (4% increase from 2008–2030 [*FAO*, 2002]). In this model, the fitness variable *RAA* mostly impacts the allocation of virtual water trade flows among different nations, and thus the shape of the node strength distribution; while the total flow parameter controls the global volume of VWT, and thus the "scale" of the distribution. Thus, the increasing total flow extrapolated from historical data implies a growth of global VWT, and the different trends of national *RAA* lead to the larger tail and heterogeneity of the projected strength distributions.

4.2. Directed Network

[19] We use population projections [*FAO*, 2010c], as well as previously used rainfall and agricultural area projections to run scenarios predicting the distributions of directed strength in 2030. We build the fitness function $\vec{q}\left(y_i^T, z_j^T\right) = \gamma^T y_i^T z_j^T$ where y^T and z^T are the projection of normalized *RAA* and population, respectively. The parameter γ^T is determined as in section 3.2., using the estimated future total flow F^{2030} = 1, 700 km^3 (vs. F^{2008} = 671 km^3). Both the distribution of national virtual water exports (s_{out}) and imports (s_{in}) are expected to increase in mean and variance (Figures 3c and 3d). In particular, the largest national virtual water export is projected to increase from about 18 km^3 in 2008 to 45 km^3 in 2030, and the largest import volume would increase from

about 18 km^3 in 2008 to 60 km^3 in 2030. Both distributions are expected to become more heterogeneous, signifying that the importance of major virtual water trade hubs will increase relative to other countries.

[20] The distribution of national import volumes is predicted to become even more heterogeneous than the distribution of exports. Indeed, the large tail of the predicted distribution of s_{in} relative to that of s_{out} shows that a few countries will concentrate most of the global import volume of virtual water, in an even larger extent than for exports. This may be explained by the fact that global population is predicted to increase, while global *RAA* is likely to decrease, and the fitness variables assigned to exporters and importers are based on the nations' *RAA* and population, respectively. Global population is predicted to increase by 23.5% between 2008 and 2030 [*FAO*, 2010c], with most of this growth taking place in emerging economies, increasingly important players in the international food trade. The population variable has proven effective to reproduce directed water flows in the past decades, illustrating the case of emerging economies with fast growing population and international food imports, like China and India. However, it may become less relevant in the future, as the population of Sub-Saharan nations will boom, but their purchasing power will likely remain low.

5. Conclusions

[21] In this Letter, we have shown that our fitness model reproduces well the topological and weighted properties of the undirected and directed virtual water trade network associated with staple food trade over a time-period of significant changes (1986–2008), assuming as sole controls each nation's GDP, rainfall on agricultural area, and population. The population variable is shown for the first time to reproduce the crucial directed flows of virtual water over time.

[22] The simplicity of these input variables imply that future scenarios may be easily run to predict the future global structure of the GVWTN. The tendency for a few countries to concentrate most virtual water trade flows has been observed in the past, in particular with the emergence of China as a major importer [*Dalin et al.*, 2012]. We found that the distribution of network flows is expected to become increasingly heterogeneous, and, thanks to our directed network modeling framework, that a few importing countries are likely to concentrate a significant portion of virtual water trade through food commodities.

[23] **Acknowledgments.** The funding support of the School of Engineering and Applied Sciences of Princeton University is gratefully acknowledged.

[24] The Editor thanks the two anonymous reviewers for their assistance in evaluating this paper.

References

Carr, J. A., P. D'Odorico, F. Laio, and L. Ridolfi (2012), On the temporal variability of the virtual water network, *Geophys. Res. Lett.*, *39*, L06404, doi:10.1029/2012GL051247.

Dalin, C., M. Konar, N. Hanasaki, A. Rinaldo, and I. Rodriguez-Iturbe (2012), Evolution of the global virtual water trade network, *Proc. Natl. Acad. Sci.*, *109*(16), 5989–5994, doi:10.1073/pnas.1203176109.

de Fraiture, C., X. M. Cai, U. Amarasinghe, M. Rosegrant, and D. Molden (2004), Does international cereal trade save water? The impact of virtual water trade on global water use, *IWMI Res. Rep. H035342*, Int. Water Manage. Inst., Battaramulla, Sri Lanka.

D'Odorico, P., J. A. Carr, F. Laio, and L. Ridolfi (2012), Spatial organization and drivers of the virtual water trade: A community-structure analysis, *Environ. Res. Lett.*, *7*, 034007, doi:10.1088/1748-9326/7/3/034007.

FAO Aquastat (2010), Information System on Water and Agriculture, http://www.fao.org/nr/water/aquastat/main/index.stm, Rome.

Field, C., et al. (2012), *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, Cambridge Univ. Press, Cambridge, U. K.

Food and Agriculture Organization of the United Nations (FAO) (2002), World agriculture: Towards 2015/2030, summary report, Rome.

Food and Agriculture Organization of the United Nations (FAO) (2010a), Faostat Detailed Trade Matrices, http://faostat.fao.org/site/291/default.aspx, Rome.

Food and Agriculture Organization of the United Nations (FAO) (2010b), Faostat Food Supply, http://faostat3.fao.org/home/index.html, Rome.

Food and Agriculture Organization of the United Nations (FAO) (2010c), Faostat Database, http://faostat3.fao.org/home/index.html, Rome.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, N. Shirakawa, Y. Shen, and K. Tanaka (2008), An integrated model for the assessment of global water resources—Part 1: Model description and input meteorological forcing, *Hydrol. Earth Syst. Sci.*, *12*, 1007–1025.

Intergovernmental Panel on Climate Change (2007), *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by M. L. Parry et al., Cambridge Univ. Press, Cambridge, U. K.

Keutelian, H. (1991), The Kolmogorov-Smirnov test when parameters are estimated from data, *Rep. CDF 1285*, U.S. Dep. of Energy, Washington, D. C.

Konar, M., C. Dalin, S. Suweis, N. Hanasaki, A. Rinaldo, and I. Rodriguez-Iturbe (2011), Water for food: The global virtual water trade network, *Water Resour. Res.*, *47*, W05520, doi:10.1029/2010WR010307.

Konar, M., C. Dalin, N. Hanasaki, A. Rinaldo, and I. Rodriguez-Iturbe (2012), Temporal dynamics of blue and green virtual water trade networks, *Water Resour. Res.*, *48*, W07509, doi:10.1029/2012WR011959.

Lilliefors, H. W. (1967), On the kolmogorov-smirnov test for normality with mean and variance unknown, *J. Am. Stat. Assoc.*, *62*(318), 399–402.

Meehl, G., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. B. Mitchell, R. Stouffer, and K. Taylor (2007), The WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, *88*, 1383–1394.

Suweis, S., M. Konar, C. Dalin, N. Hanasaki, A. Rinaldo, and I. Rodriguez-Iturbe (2011), Structure and controls of the global virtual water trade network, *Geophys. Res. Lett.*, *38*, L10403, doi:10.1029/2011GL046837.

The World Bank (2012), The World Bank Databank, http://data.worldbank.org/indicator, Washington, D. C.

Yang, H., L. Wang, K. Abbaspour, and A. Zehnder (2006), Virtual water trade: An assessment of water use efficiency in the international food trade, *Hydrol. Earth Syst. Sci.*, *10*, 443–454.