

Wheat trade tends to happen between countries with contrasting extreme weather stress and synchronous yield variation

Srishti Vishwakarma ¹, Xin Zhang ¹✉ & Vyacheslav Lyubchich ²

Extreme weather poses a major challenge to global food security by causing sharp drops in crop yield and supply. International crop trade can potentially alleviate such challenge by reallocating crop commodities. However, the influence of extreme weather stress and synchronous crop yield anomalies on trade linkages among countries remains unexplored. Here we use the international wheat trade network, develop two network-based covariates (i.e., difference in extreme weather stress and short-term synchrony of yield fluctuations between countries), and test specialized statistical and machine-learning methods. We find that countries with larger differences in extreme weather stress and synchronous yield variations tend to be trade partners and with higher trade volumes, even after controlling for factors conventionally implemented in international trade models (e.g., production level and trade agreement). These findings highlight the need to improve the current international trade network by considering the patterns of extreme weather stress and yield synchrony among countries.

¹Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, MD 21532, USA. ²Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, MD 20688, USA. ✉email: xin.zhang@umces.edu

Extreme weather events, such as drought, flood, and heat-wave threaten food security from regional to global scales through the resulting sharp decline in the availability, affordability, and adequate utilization of food^{1,2}. During 2003–2013, extreme weather events caused marked damage of USD 30 billion to agricultural productivity³. Crop production was impacted the most, with yield reductions^{3–7} introducing price volatility in the food systems^{1,8}, affecting food trade, the welfare of farmers, and economic development, especially in low-income or import-dependent countries^{1,9–11}.

International crop trade can potentially alleviate the negative impacts of extreme weather events on food security by exporting food commodities from surplus to deficit regions¹². Currently, international trade accounts for 23% of the global food supply for humans through major food commodities. Wheat, a crop essential for people's daily caloric and protein needs, was produced at the rate of 683 million tonnes annually and 147 million tonnes was traded globally, accounting for 22% of the crop trade (in caloric content) in 2009^{13–15}. However, a heavy reliance on imports from other countries or the global market may expose a country to the yield and market variations outside of the country's jurisdiction and consequently introduce additional risk to the country's food supply. For example, the 2010 heatwave in Russia triggered export restrictions for wheat, led to a wheat shortage and price spike in the Middle East, where over 1/3 of the wheat supply is from Russia, and potentially contributed to the destabilization of the region^{9,16}. A simultaneous drop in yields of major exporters may disrupt the global trade network and food supply, and a synchronous yield fluctuation between trade partner countries may further exacerbate the problem. Therefore, the controversial role of international trade in addressing the food security challenge is associated with patterns of extreme weather events and yield fluctuations; however, such associations remain poorly understood and require an in-depth investigation¹⁷.

The occurrence and volume of the trade between countries have been often investigated as results of comparative advantages in producing food commodities (e.g., more efficient use of water and land resources), as well as many socioeconomic factors such as geographical proximity of countries, population, agricultural productivity, language, contiguity, level of economic development, and trade agreements^{18–22}. Several recent studies have evaluated the impacts of climate factors¹⁹, but most focus on the average state of the climate, such as annual rainfall, annual evapotranspiration, and annual temperature^{19,23,24}. Many studies have examined the impacts of natural disasters^{25,26}, most of which are quantified based on deaths and losses due to rare events like flood or drought²⁷ and are not specific to the impacts of extreme weather events on crop production^{28,29}. Only a few studies investigated the impacts of extreme weather stress and synchronous crop yield fluctuations^{30–33}, but their focus was on the impacts on food price fluctuations or trade volumes for individual countries, and not on the changes in the bilateral trade network.

In addition, the investigation of drivers for trade links has been limited to statistical approaches that were not designed to handle complex network data or derive data-driven relationships. Prior research of the potential drivers used regression models that impose multiple restrictive assumptions on the shapes of relationships (e.g., linear, log-linear, or log-log) and data distribution^{24,34–38}. These regression approaches focus on modeling the attributes of network vertices without considering the structure of the network. In contrast, network analysis emphasizes both the vertices and structure of the network; thereby, it is capable of handling complex relationships^{39,40}. Only recently, the statistical exponential random graph models (ERGMs) have been used to investigate the relationships between international trade

links (or volume) and their potential drivers^{22,41,42}. The ERGM has proven to be superior in capturing zero-inflated trade volumes and complex trade patterns (e.g., third-party relations and clustering), compared to traditional approaches²². However, these recent studies still impose parametric assumptions and are not flexible enough for modeling highly nonlinear relationships. Despite the success of machine learning approaches such as random forest (RF) in handling large volumes of complex data and deriving nonlinear relationships from the data, such data-driven approaches have been rarely utilized in trade analysis⁴³.

To address these knowledge gaps, we proposed network-based covariates for studying the international trade network of wheat during the period from 2005 to 2014 using modern statistical and machine learning models. In addition to commonly used geopolitical (e.g., contiguity) and economic (e.g., regional trade agreements) factors, two network-based covariates were developed to characterize the extreme weather stress and yield synchrony, namely the difference in extreme weather stress (DEWS) and short-term synchrony (STS) of crop yield anomalies between countries. The DEWS network, derived using principal components of weather stress indices (see Methods), quantifies by how much the weather stress in an exporting country is stronger than the stress in an importing country. The STS network is represented by the correlation of fluctuations in crop yield, where the fluctuations are defined as deviations from the current average yield. The STS allows us to quantify how strongly the deviations of yields in one country correspond to similar deviations in other countries. A positive STS indicates synchronous fluctuations, while a negative STS indicates asynchronous fluctuations. Using the two network-based covariates, we hypothesize that 1) Higher weather stress difference (i.e., higher DEWS) leads to higher import volumes and likelihood of trade partnerships, and 2) Countries with synchronized yield anomalies (i.e., positive STS) are less likely to trade or have lower trade volumes.

To accommodate the complexity and network structure of the data, we applied ERGM and RF to model trade linkages and volume between countries (see Methods section for details). The ERGM was selected as the most suitable statistical method, and RF, being one of the best-performing machine learning techniques, was selected as a nonparametric alternative for benchmarking the results. We compared the performance of ERGM and RF in cross-validation for different types of networks, which highlighted the strengths and weaknesses of these alternative methods.

The analysis reveals that the current wheat trade network tends to have partnerships between countries with more synchronized yield fluctuation, exposing countries to synchronized yield failure. In addition, our models show that countries with larger differences in extreme weather stress tend to have higher import volumes and more trade partnerships. Therefore, this study suggests the need to consider extreme weather stress and yield synchrony in the design of trade networks to improve the stability and fairness of the global food system.

Results

Extreme weather stress for wheat production. Cold and heat stresses were identified as the major contributors to the variability of extreme weather indices developed for a country's wheat production. A total of 17 indices were used to quantify weather stresses (including heat stress, cold stress, flood, and drought) during the growing period for wheat in 115 countries for the years 2005–2014 (see Methods and Supplementary Note 3). The first two principal components of the 17 indices, dominated by cold and heat stress, represent 65% and 22.7% variance of the weather index matrix, respectively (Supplementary Fig. 2).

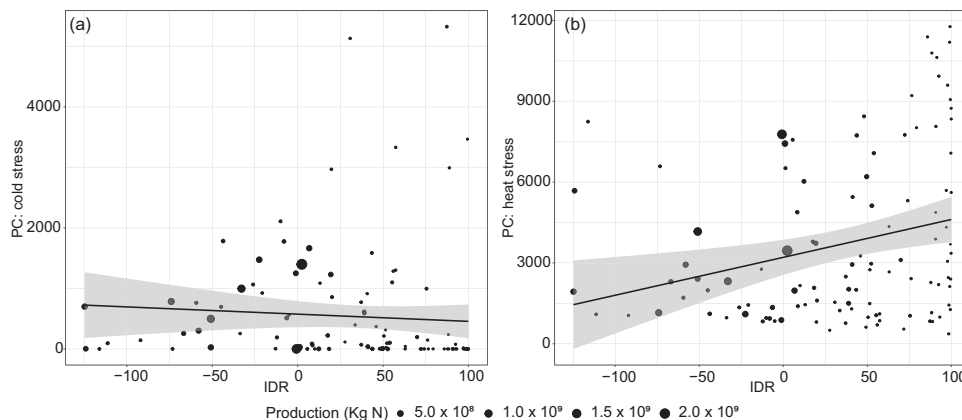


Fig. 1 Relationship between import dependency ratio and weather stress. Relationships between the 2005–2014 import dependency ratio (IDR; Eq. 1) and derived principal components (PCs) representing the weather stress: **(a)** cold stress, **(b)** heat stress. Positive IDR means higher import dependency, while a negative IDR means that a country is a net exporter. Each black point represents a country, the size of the point corresponds to the average wheat production during 2005–2014. The black lines represent the estimated linear relationships between weather stress and IDR (p -value = 0.460 for cold stress and 0.005 for heat stress), gray shaded areas correspond to 95% confidence intervals.

The dominant principal components of the extreme weather indices are not significantly correlated with production levels across countries, while the heat stress indices are correlated with the import dependency (Fig. 1; detailed results in Supplementary Table 3 and Supplementary Figs. 13–29). It suggests that the scale of wheat production in a country was not necessarily affected by the extreme weather stress in the wheat-producing region, but a country's dependency on wheat imports was associated with higher heat stress. Furthermore, out of 115 major wheat-producing countries in our study, 56 countries have an import dependency ratio (IDR) greater than 50% (mostly in developing countries), while only 27 countries are net exporters. Since around half of the countries are highly import-dependent for wheat, trade plays an important role in ensuring a stable wheat supply. The pairwise relationships between weather stress and other major characteristics of trade (such as the number of linkages and trade volume) are similar: the countries facing higher heat stress (or lower cold stress) are likely to have fewer trade partners for exports; and countries with higher cold stress tend to have more import trade partners (Supplementary Figs. 8–11 and Supplementary Table 3).

Relationships between trade networks and extreme weather stress.

Using the principal components of extreme weather indices as part of the covariates, we modeled the bilateral trade networks (one weighted by trade volume and one without the weights) of wheat with ERGM and RF model separately. The two models observe similar general relationships between trade networks and their potential drivers, but the performances of the two models vary. To evaluate the performance of each model, we conducted cross-validation. The results show that ERGM, with an error rate of 5.17%, was more accurate than RF in predicting trade presence/absence (i.e., the trade network without weight by trade volume), while RF was more accurate in predicting trade volume (Tables 1 and 2; see the distribution of cross-validated errors in Supplementary Figs. 5–6). Hence, throughout the rest of the paper, we report the modeling results for trade linkages and trade volumes based on ERGM and RF, respectively.

Modeling results from both ERGM and RF show that country pairs with larger differences in the levels of extreme weather stresses are more likely to be trade partners. The ERGM shows that more severe heat stress in the importing country compared to an exporting country (i.e., $DEWS_{heat} < 0$) corresponds to a higher likelihood of trade link formation. Vice versa, trade

partnerships are less likely if the exporting country is experiencing larger heat stress than the importer does (i.e., $DEWS_{heat} > 0$; Table 1, and Supplementary Fig. 12b). These model results align with the observed relationship between import dependency and heat stress (Fig. 1b).

The differences of both heat and cold stress between countries have significant relationships with trade volume. The RF shows overall higher trade volumes correspond to higher heat stress in importing countries (i.e., when $DEWS_{heat} < 0$, compared with $DEWS_{heat} > 0$, similar to the ERGM results); however, the relationship is not exactly linear, and the trade volumes increase marginally for $DEWS_{heat}$ around zero (Fig. 2b). Higher trade volume is predicted when differences in cold stress between partners exist (i.e., $DEWS_{cold} \neq 0$; Fig. 2a); however, in contrast to the heat stress, two upper deciles of $DEWS_{cold}$ are associated with higher trade volumes. In particular, the biggest spike in Fig. 2a is driven by Germany, i.e., a large exporter that may experience more severe cold stress than its trade partners do ($DEWS_{cold} > 0$). The cases of $DEWS_{cold} > 3000$ are dominated by Japan, Mongolia, and South Korea in the exporter role; hence, the corresponding average trade volumes decline from the peak values (Fig. 2a). The RF rankings of variables by their predictive importance are consistent between the weighted and unweighted trade networks. The economic and difference in production covariates are ranked as the top predictors followed by the STS and DEWS. Among the other factors, common official language and contiguity are ranked eighth and ninth, while regional trade agreements (RTA) is ranked tenth (Supplementary Figs. 32 and 33). Permutation-based assessment of statistical significance of the importance values identified all the variables in the RFs as statistically significant, which means the RF models have one additional variable ($DEWS_{cold}$) compared with the ERGMs for the unweighted directed trade network.

The role of yield synchrony. In addition to extreme weather stress, the STS of crop yield anomalies also demonstrates a significant relationship with the wheat trade networks regarding the presence/absence of trade links and trade volumes. More specifically, the ERGM for the unweighted network shows that STS is positively associated with the likelihood of trade partnerships (Table 1). In the weighted trade network, RF detects a nonlinear relationship characterized by the overall accelerating increase of trade volume with the increase in STS (main body of the distribution; Fig. 2c). However, the first decile of STS, comprising

Table 1 Summaries of the models for 2005–2014 unweighted directed (trade presence) international wheat trade network.

Model	Covariate	Coefficient	p-value	Error		
ERGM	DEWS _{heat}	-2.45×10^{-5} (4.39×10^{-6})	<10 ⁻⁴	5.17%		
	STS	0.13 (4.40×10^{-2})	0.0036			
	Production	7.02×10^{-10} (4.91×10^{-11})	<10 ⁻⁴			
	GDP	1.91×10^{-13} (1.28×10^{-14})	<10 ⁻⁴			
	GDP per capita	6.36×10^{-6} (6.45×10^{-7})	<10 ⁻⁴			
	GATT/WTO	-0.455 (3.52×10^{-2})	<10 ⁻⁴			
	RTA	0.97 (0.102)	<10 ⁻⁴			
	Distance	-2.82×10^{-5} (2.16×10^{-6})	<10 ⁻⁴			
	Contiguity	2.30 (0.167)	<10 ⁻⁴			
	Common official language	0.27 (6.13×10^{-2})	<10 ⁻⁴			
	RF	DEWS _{heat} , DEWS _{cold}	Supplementary Fig. 4		0.01, 0.01	21.15%
		STS			0.02	
		Production			0.01	
GDP			0.01			
GDP per capita			0.01			
GATT/WTO			0.01			
RTA			0.01			
Distance			0.01			
Contiguity			0.01			
Common official language			0.01			

ERGM exponential random graph model, RF random forest, DEWS difference in extreme weather stress, STS short-term synchrony, GDP gross domestic product, GATT The General Agreement on Tariffs and Trade, WTO World Trade Organization, and RTA Regional Trade Agreements. Standard errors of the coefficients are shown in parentheses. Errors are the average cross-validated misclassification errors for trade presence.

Table 2 Summaries of the models for 2005–2014 weighted directed (trade volume) international wheat trade network.

Model	Covariate	Coefficient	p-value	Error		
ERGM	DEWS _{cold}	4.82×10^{-6} (7.55×10^{-7})	<10 ⁻⁴	1.48		
	DEWS _{heat}	-4.59×10^{-6} (2.51×10^{-7})	<10 ⁻⁴			
	STS	2.14×10^{-2} (3.97×10^{-9})	<10 ⁻⁴			
	Production	2.23×10^{-10} (9.54×10^{-12})	<10 ⁻⁴			
	log10(GDP)	-1.43×10^{-2} (2.06×10^{-9})	<10 ⁻⁴			
	GDP per capita	1.41×10^{-6} (2.68×10^{-8})	<10 ⁻⁴			
	GATT/WTO	-9.79×10^{-2} (1.18×10^{-8})	<10 ⁻⁴			
	RTA	0.186 (1.30×10^{-8})	<10 ⁻⁴			
	Distance	-5.95×10^{-6} (9.22×10^{-8})	<10 ⁻⁴			
	Contiguity	0.49 (2.19×10^{-8})	<10 ⁻⁴			
	Common official language	3.56×10^{-2} (2.48×10^{-8})	<10 ⁻⁴			
	RF	DEWS _{heat} , DEWS _{cold}	Figure 2		0.01, 0.01	0.77
		STS			0.03	
Production			0.01			
log10(GDP)			0.01			
GDP per capita			0.01			
GATT/WTO			0.01			
RTA			0.01			
Distance			0.01			
Contiguity			0.01			
Common official language			0.01			

ERGM exponential random graph model, RF random forest, DEWS difference in extreme weather stress, STS short-term synchrony, GDP gross domestic product, GATT The General Agreement on Tariffs and Trade, WTO World Trade Organization, and RTA Regional Trade Agreements. Standard errors of the coefficients are shown in parentheses. Errors are the average cross-validated mixed errors for trade volume.

the most asynchronous pairs of countries, is also characterized by a spike in trade volume (Fig. 2c). This illustrates that countries with perfect asynchrony (STS ≈ -1) and synchrony (STS ≈ 1) of yield fluctuations tend to trade more.

The role of other factors. All our ERGM and RF models also include the following covariates that have been considered as important for the formation of trade linkages: population-weighted distance, contiguity, common official language, gross domestic product (GDP), GDP per capita, General Agreement on

Tariffs and Trade (GATT) /World Trade Organization (WTO) memberships, RTA, and difference in production between countries. The modeling results further confirmed the important role of these factors. The ERGM results show that trade partnerships are more likely to occur between countries that are closer to each other, contiguous, or have a common official language (Table 1), which aligns well with the existing findings in the literature. The RF results show, similarly to ERGM, higher trade volumes for countries that are contiguous and have a common official language (Fig. 2e, f), and an overall negative relationship between trade volume and distance (Fig. 2d). However, RF was

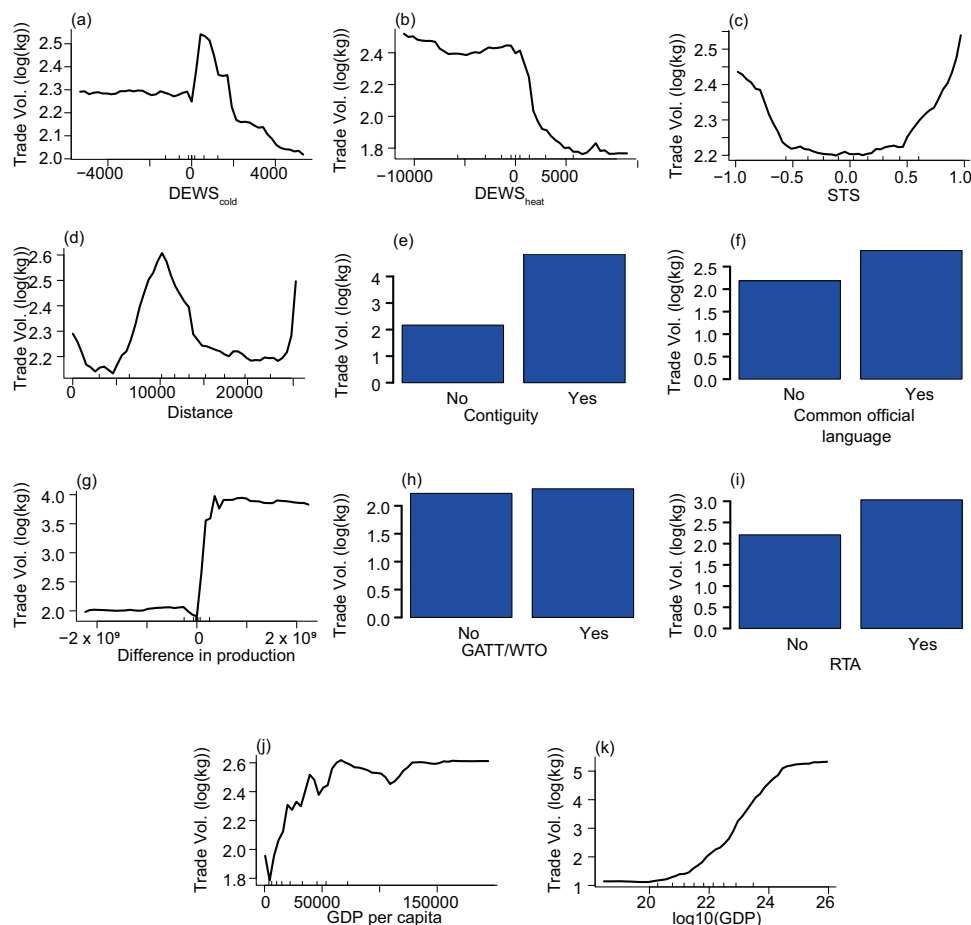


Fig. 2 Random forest partial dependence plots for trade volume in 2005–2014. The x-axes represent the considered covariates: (a) difference in extreme cold stress ($DEWS_{cold}$), (b) difference in extreme heat stress ($DEWS_{heat}$), (c) short-term synchrony (STS), (d) distance, (e) contiguity, (f) common official language, (g) difference in production, (h) The General Agreement on Tariffs and Trade (GATT)/ World Trade Organization (WTO), (i) Regional Trade Agreements (RTA), (j) gross domestic product (GDP) per capita, and (k) GDP. The inner tickmarks on the x-axes represent deciles of the variables. The y-axis represents the marginal effect of the covariate on wheat trade volume. The blue bars show marginal effects caused by categorical covariates, while black lines show marginal effects of continuous covariates.

also able to model nonlinearity in the latter relationship, characterized by substantial spikes in trade volume around the deciles 3–4 and 9–10 of the population-weighted distance (Fig. 2d).

The inclusion/exclusion of these covariates in the ERGM and RF models does not affect the above results regarding the relationships between trade networks and extreme weather stresses, as well as yield synchrony, further confirming the robustness of the modeling results. For example, countries closer to each other tend to have more synchronized yield; however, the ERGM results show a significant positive association of trade partnerships with STS regardless of whether the distance variable is included or excluded (Supplementary Table 4). This test suggests that the positive relationship between trade networks and STS is not only due to the positive relationship between STS and distance but could be an outcome of other factors that are not included in the models (e.g., cultivars, and technology and management practices in agriculture).

The difference in production level between exporting and importing countries shows a significant positive association with the international wheat trade network (Fig. 2g and Supplementary Fig. 12a). The amount of production plays a critical role in a country's decision to engage in international trade. The modeling results from both ERGM and RF show that the likelihood of trade partnership is higher when country pairs have larger differences

in their production levels. The ERGM results show a higher likelihood of trade partnership when production is higher in exporting country (i.e., difference in production > 0) compared to when it is higher in importing country (i.e., difference in production < 0).

In RF, a non-linear relationship between difference in production and trade volume is observed (Fig. 2g). A substantial increase in trade volume is predicted when the production is higher in exporting country (i.e., difference in production > 0) compared with difference in production < 0 , such a pattern is similar to ERGM results. The cases with difference in production $> 10^9$ (i.e., 1 million tonnes; upper decile) mostly include the Russian Federation, the USA, and India as the major exporters. Specifically, the highest marginal increase in trade volume (Fig. 2g) is driven by the Russian Federation, the USA, France, and Canada highlighting their major contribution to the international wheat trade.

In addition to the above covariates, we tested several other economic factors, including GDP, GDP per capita, GATT/WTO membership, and RTA. These variables are frequently used to account for multilateral resistance, total expenditure, and market size in international trade^{44,45}. The ERGM results show that trade partnerships are more likely to occur for countries with higher GDP/GDP per capita and involvement in regional trade agreements, while the possibility of trade partnership reduces

when countries are in GATT/WTO membership. In the RF, the GATT/WTO shows a similar marginal increase in volume of trade for both member and non-member countries, indicating that the imports and exports are occurring between countries irrespective of their trade membership. However, including these factors in the models does not change the relationship between the trade network and the two main covariates (i.e., DEWS and STS).

Discussion

While few trade partnerships have been established with direct consideration of extreme weather stress and yield synchrony among countries, significant relationships between these factors and the wheat trade network have been revealed by our investigation into the historical records for the period of 2005–2014. The revealed relationships have strong implications for the resilience of wheat trade networks and food supply for countries around the world.

The significant positive associations between trade linkages and synchronous yield fluctuations indicate that in the current wheat trade network, trade partnerships tend to be established between countries with more synchronized yields. Therefore, the current wheat trade network is potentially vulnerable to synchronous yield failure. A recent study projected that synchronous failure in wheat yield can lead to about a 33% reduction in wheat production⁴⁶, which is equivalent to half of the projected wheat consumption by 2050⁴⁷. Hence, the synchronous yield failure can lead to dramatic disruptions to wheat supply for a country and the world, and consequently threaten food security in countries, especially those low- and middle-income countries that are dependent on imports^{1,11}.

Higher likelihood of trade partnerships and higher import volume are found to be associated with higher differences in extreme weather stress when importing countries have higher stress than their exporting partners. This finding indicates the role of international trade in alleviating the long-term production stress caused by frequent extreme weather. Under the possible higher weather stress due to future climate change, the wheat trade network is expected to change, and increases in import linkages are likely to occur, particularly in tropical countries that are already import-dependent. Consequently, whether these countries can establish more trade linkages on time to combat the climate-change-induced stress in domestic production and stabilize their wheat supply is critical for food security in these countries.

Our analysis demonstrates the need to factor in the consideration of extreme weather stress and yield synchrony when designing trade agreements and establishing trade partnerships. Countries have been choosing their trade partner mainly based on economic benefits (e.g., promising duty- and quota-free access to the international market) or geopolitical relations, while few consider the extreme weather stress and yield variation in their potential partner countries. While the partnership is economically beneficial for the countries involved, it may lead some countries to become highly dependent on wheat supply from a few countries or regions, and potentially expose these countries to a higher risk of food insecurity when trade partners are facing synchronous yield losses. For instance, Ukraine's substantial (~77%) wheat production losses in 2007 compared to 2006 forced its trade partners to buy wheat from other countries (e.g., Australia, France, the USA, Russia, Kazakhstan) and relax import barriers to meet their domestic food demand⁴⁸. Such rapid alterations in trade partners within a year overwhelm not only the major exporters through reduced stock-to-use ratio and doubled export

volume in a short period, but also importers who have been relying on the supply from Ukraine or the global market, potentially contributing to spikes in wheat prices^{16,48}. Importers from least-developed nations are particularly vulnerable to such rapid changes in the global market, as they usually have less political and economic capacity to establish new partnerships or to adapt to higher market prices. Therefore, the economic benefits and abundant food supply, brought by the trade partnerships, can potentially be compromised by the occurrence of extreme weather stress and synchronized yield failure. To mitigate these potential negative outcomes, countries, as well as multilateral trade agreements, need to carefully consider their portfolio of trade partners by 1) identifying whether their trade partners have synchronized yield patterns; 2) determining the level of extreme weather stress in each partner and the impact of climate change. To ensure a more stable crop supply, a country or a trade agreement should favor trade partners with asynchronized yield patterns and less extreme weather stress, in addition to economic benefits. In addition, the examination of yield synchrony and extreme weather stress in existing partnerships will inform the design of mitigation measures such as crop reserves to minimize the disruption of synchronized yield failure.

We tested two models, ERGM and RF, for investigating the trade linkages and trade volume. The goal of testing the two models is to assess the linear (via ERGM) and nonlinear (via RF) relationships between trade networks and drivers. While the two modeling approaches provide similar results on the general relationships between the trade networks and a range of covariates, we find that ERGM performed better when investigating the unweighted network, and RF performed better for the network with trade volume (Tables 1 and 2). The reason for the better cross-validation performance of ERGM in an unweighted network but not in a weighted network is the relative simplicity and thus robustness of the model when capturing the presence/absence information and, at the same time, lack of flexibility when modeling the actual trade volumes. In contrast, RF automatically captures the nonlinearity of the relationships but tends to overfit the data; hence, the results can be less robust, such as in the case of the unweighted network.

RF performed better than ERGM on the weighted network in the cross-validation study, which implies that the nonlinear relationships the RF was able to capture were important in predicting trade volume. However, the difference in errors between models is substantial. Although RF has lower errors, it still needs improvements in providing robust predictions of trade volume (see Table 2 and Supplementary Figs. 5 and 6). The correlation of observed and predicted percentage contributions of each country in global export volume is 0.91 (Supplementary Fig. 34). However, the correlation between observed and predicted trade volumes is not as strong (Supplementary Fig. 35). The performance can potentially be improved by considering in the modeling process other commonly used drivers, the impact of physical limits of crop production on trade volumes, and adjustments for imbalanced (zero-inflated, in case of numeric response such as trade volume) distributions. Substantial progress has been made in RF classification to deal with imbalanced datasets^{49–52}, but developments for RF regression with zero-inflated data are still ongoing⁵³. Even with the current limitations and uncertainties in ERGM and RF, the models can establish a reliable general relationship between trade networks (with and without the weight of trade volume) and extreme weather stress and yield synchrony. Further research is needed to improve the modeling and prediction of trade volume in the context of climate change by incorporating the production and demand levels of both the importing and exporting countries.

Conclusions

Our analysis suggests that the two factors, the level of extreme weather stress and synchrony of crop yield fluctuations, substantially affect the international wheat trade network. Country pairs with larger differences in heat stress are more likely to have trade connections and higher trade volumes. Meanwhile, in the current wheat trade network, trade partnerships are more likely to be established between countries with synchronized yield fluctuations. This represents a systemic risk in the current global wheat market since synchronized yield failure can disrupt wheat supply and intensify food insecurity for both partnering countries. Other fundamental drivers in the analysis (i.e., production level, and economic and geopolitical factors) show significant relationships with trade link and volume, further confirming their importance in the international trade network. Our results demonstrate the need to consider extreme weather stress and yield synchrony in the trade policy framework to improve the stability and fairness of the global food system.

Methods

Weather, wheat yield, and trade data. Major datasets used in this study include daily weather records⁵⁴, annual wheat yield⁵⁵, annual wheat production⁵⁵, and bilateral trade volume¹⁵ for the period of 2005–2014. To assess crop-specific weather conditions for countries around the world, global maps of crop calendar⁵⁶ and harvested area⁵⁷ were used. The other datasets used in this study include the distance between countries⁵⁸, and their official languages⁵⁹, contiguity⁵⁸, gross domestic product (GDP)⁶⁰, GDP per capita⁶⁰, General Agreement on Tariffs and Trade (GATT)/World Trade Organization (WTO) memberships, and regional trade agreements (RTA)^{61,62} (Supplementary Table 1). The reason to include these additional datasets is to incorporate factors that are considered important for trade connections. Countries with at least seven years of yield data in the analyzed period were selected for this study (115 countries, see Supplementary Note 1). All data used in this study are publicly available from the corresponding sources.

Import dependency ratio. Import dependency ratio (IDR) for each country was calculated as the ratio of net imports (i.e., *imports*—*exports*) and total production of the crop (kg) and net imports (kg):

$$IDR = \frac{\text{net imports}}{\text{production} + \text{net imports}}. \quad (1)$$

Network-based covariates for extreme weather stress. To assess the extreme weather stress in wheat-producing countries and the relationship among countries, we developed the Difference in Extreme Weather Stress (DEWS) covariate, following four steps outlined below. DEWS was used to address the hypothesis that trade is more likely between a country with low weather stress (country capable of producing high yields due to good weather) and a country with high weather stress (country willing to buy because of yield losses due to bad weather), i.e., in the cases with large DEWS. If both trade partners experience good or bad weather for growing the crop, DEWS is low.

First, we comprehensively characterized local agricultural weather conditions. We calculated 17 agricultural weather indices⁶³ (Supplementary Table 2) for each spatial grid cell during the growing season of wheat with crop calendar in each year from 2005 to 2014 (Supplementary Algorithm 1 and Supplementary Note 2). The weather indices represent extreme weather stress as they quantify the temperature and precipitation conditions that are outside the optimal growing conditions (i.e., comfort zone) of wheat⁶³. Second, we aggregated the weather indices for relevant areas. The gridded maps of weather indices from the first step were aggregated for each country with the spatial grid weights proportional to the harvested area for wheat (e.g., weather index in grid cells without wheat production was disregarded) (Supplementary Algorithm 1, Supplementary Note 2, and Supplementary Fig. 36). This resulted in 17 weather indices for 115 countries for 10 years. Third, we reduced the dimensionality of the weather data to provide a combined assessment of the weather conditions and to avoid potential problems of multicollinearity and overfitting in later modeling. In the principal component analysis (PCA) of the extreme weather indices (see Supplementary Note 4), the first two principal components (PCs) account for 87.7% of the variation, and they represent cold and heat extreme weather stress conditions, respectively, in a country's wheat-producing region (Supplementary Fig. 2). Fourth, we defined DEWS for cold stress (DEWS_{cold}) (or heat stress; DEWS_{heat}) as the difference in the cold stress PC (or heat stress PC) in two countries, where DEWS < 0 means higher stress in importing country, and DEWS > 0 means higher stress in exporting country.

Network-based covariate for yield synchrony. To assess associations of synchrony of crop yield fluctuations with trade, we developed the covariate short-term

synchrony (STS). STS was defined as the correlation of yield fluctuations between countries. The yield fluctuations for each country were obtained by removing trends from the yield time series, where the trend was estimated using a locally weighted regression approach (see Supplementary Note 5 and Supplementary Fig. 3).

Trade networks. Based on the bilateral wheat trade flows during 2005–2014, two types of trade networks were developed and examined in this study, namely unweighted directed and weighted directed networks. Both networks consist of nodes and edges, where nodes represent countries and edges represent trade connections between countries. The edges are directed, representing the direction of trade flows from one node to another. The unweighted directed network does not assign a weight to the edges, while the weighted directed network assigns a weight for each edge based on the trade volume (measured in kg). Between 2005 and 2014, the international trade network included 115 nodes (i.e., countries) and 5726 directed edges, resulting in an average in/out-degree of 50 (Supplementary Figs. 7, 30, and 31).

Exponential random graph model. The exponential random graph model (ERGM) was chosen as the most suitable and powerful statistical approach for network modeling. ERGM is an exponential probability model used to predict edges (or weights on edges) between nodes in a network Y^{64} :

$$\Pr_{\theta, \mathcal{Y}}(Y = y) = \frac{\exp\{\theta^T g(y, X)\}}{\kappa(\theta, y)}, y \in \mathcal{Y}, \quad (2)$$

where $\theta \in \mathbb{R}$ represents model coefficients, $g(y, X)$ are statistics corresponding to y and covariates X , $\kappa(\theta, y)$ is the normalizing factor, and \mathcal{Y} is the set of all possible networks for the given number of nodes.

In this study, the network Y represents network realization of the observed network y , covariates X include the DEWS, STS, distance between countries, contiguity, common official language, GDP, GDP per capita, RTA, GATT/WTO membership, and difference in production. The ERGM is tested for both the trade networks (i.e., weighted directed and unweighted directed). See Supplementary Note 6 for details on the model specification in R software. In addition to the above-defined ERGM, several different combinations of ERGM are tested (see Supplementary Note 15 and Note 18).

The sign and magnitude of the ERGM coefficients for each covariate (Tables 1 and 2) demonstrate the likelihood of link formation in the trade network²². A significant (p -value < 0.05) positive coefficient shows that higher values of the covariate are associated with a higher likelihood of the link formation, while a negative coefficient corresponds to a lower likelihood. Statistical significance of the coefficients is assessed using t -test at 0.05 significance level.

Random forest model. The random forest (RF) model⁶⁵ was used as an alternative nonparametric method for modeling the two trade networks. RF is one of the best-performing machine learning techniques, hence serves as a valuable benchmark for validating the results. Compared with other popular machine learning models, such as neural networks, RF has fewer tuning parameters. Some of the important tuning parameters in RF are the number of variables to examine at each split and the minimal size of terminal nodes, both of which have sensible defaults⁶⁶. In contrast, neural networks rely on many more hyperparameters, many of which do not have default values and have to be set by the user (for example, number of layers and number of units in each layer, dropout rates, size and number of filters in convolutional layers)⁶⁷. We chose RF due to its competitive performance, ease of specification, and rising popularity in agricultural studies^{68–70}.

Unlike ERGM, random forest is a machine learning tool that investigates relationships between covariates and can model nonlinear and combined effects of several covariates automatically. RF is a set of regression trees, where each tree is trained on a bootstrap sample of the original data, and only a random subset of covariates is assessed at each tree split (see details of RF implementation in Supplementary Note 7). The relationships captured by RF are investigated through partial dependence plots that show marginal changes in the response variable due to the changes in the covariate, while all other covariates are kept at their observed levels. In addition to the partial dependence plots, we assessed the importance of the covariates based on the mean decrease in the predictive accuracy of a regression tree when values of one of the covariates are randomly permuted⁶⁶. We further ranked the covariates based on their importance and assessed their statistical significance (at 0.05 level) using the Altman permutation algorithm with 100 permutations⁷¹.

Model assessment. Validation and benchmarking are essential steps for evaluating a model's performance and how it compares to competitive alternatives. We evaluated the model performance through cross-validation. In the cross-validation, five years of data from the observed period were randomly selected for training the models, and the remaining five years were used to obtain predictions and calculate the errors of the model predictions (see details in Supplementary Note 8).

Two types of errors were calculated in the cross-validation. For the models of the unweighted network, we used misclassification error (MCE):

$$MCE = 1 - n^{-1}(TP + TN), \quad (3)$$

where n is the sample size, TN and TP are the counts of true positive and true negative classifications.

For the models of the weighted network, we calculated mixed error (ME):

$$ME = n^{-1} \sum_{i=1}^n \frac{|OT_i - PT_i|}{1 + |OT_i|}, \quad (4)$$

where OT_i is the observed trade volume, and PT_i is the predicted trade volume for i -th trade connection. The cross-validation was carried out 100 times and the two types of errors were calculated each time.

Data availability

The weather, harvested area, and crop calendar maps are obtained from ECMWF (<https://www.ecmwf.int/>), Earthstat (<http://www.earthstat.org/>), and Center for Sustainability and the Global Environment (<https://sage.nelson.wisc.edu/data-and-models/datasets/crop-calendar-dataset/>), respectively. The wheat yield and production data are from Zhang et al. (2015) (<https://www.nature.com/articles/nature15743>). The bilateral trade, per capita GDP and GDP data are accessed through the World Bank (<https://data.worldbank.org/>). The datasets of General Agreement on Tariffs and Trade (GATT), World Trade Organization (WTO), Regional Trade Agreements (RTA), contiguity, common official language, and distance are from the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) (<http://www.cepii.fr/>). The data used for generating the figures are available in Zenodo (<https://doi.org/10.5281/zenodo.7130929>).

Code availability

The R-scripts generated during the study are available on Zenodo (<https://doi.org/10.5281/zenodo.7130929>).

Received: 2 August 2022; Accepted: 14 October 2022;

Published online: 01 November 2022

References

1. FAO, IFAD, UNICEF, WFP & WHO. The State of Food Security and Nutrition in the World 2019. Safeguarding against economic slowdowns and downturns. (2019).
2. Gaupp, F., Pflug, G., Hochrainer-Stigler, S., Hall, J. & Dadson, S. Dependency of Crop Production between Global Breadbaskets: A Copula Approach for the Assessment of Global and Regional Risk Pools. *Risk Anal* **37**, 2212–2228 (2017).
3. FAO. Impact of Natural Hazards and Disasters on Agriculture and food security and nutrition: a Call for Action to build resilient livelihoods. (2015).
4. Myers, S. S. et al. Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition. *Annu. Rev. Public Health* **38**, 259–277 (2017).
5. Rosenzweig, C. et al. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci.* **111**, 3268 LP–3263273 (2014).
6. Butler, E. E., Mueller, N. D. & Huybers, P. Peculiarly pleasant weather for US maize. *Proc. Natl. Acad. Sci.* **115**, 11935–11940 (2018).
7. Lobell, D. B., Schlenker, W. & Costa-Roberts, J. Climate Trends and Global Crop Production Since 1980. *Science* **333**, 616–620 (2011).
8. Baldos, U. L. C. & Hertel, T. W. The role of international trade in managing food security risks from climate change. *Food Secur* **7**, 275–290 (2015).
9. Mitra, S. & Josling, T. Agricultural Export Restrictions: Welfare Implications and Trade Disciplines. *International Food and Agricultural Trade Policy Council (IPC)* (2009).
10. Dellink, R., Hwang, H., Lanzi, E. & Chateau, J. International trade consequences of climate change. OECD Trade and Environment Working Papers (2017). <https://doi.org/10.1787/9f446180-en>.
11. Puma, M. J., Bose, S., Chon, S. Y. & Cook, B. I. Assessing the evolving fragility of the global food system. *Environ. Res. Lett.* **10**, 024007 (2015).
12. Anderson, W. B., Seager, R., Baethgen, W., Cane, M. & You, L. Synchronous crop failures and climate-forced production variability. *Sci. Adv.* **5**, 1–10 (2019).
13. Odorico, P. D., Carr, J. A., Laio, F., Ridolfi, L. & Vandoni, S. Feeding humanity through global food trade. *Earth's Futur* **2**, 458–469 (2014).
14. FAOSTAT. FAOSTAT: Statistical database of production quantity from Crops and livestock products. FAO, Rome, Italy. Retrieved: 30 Nov 2021. (2021). Available at: <https://www.fao.org/faostat/en/#data/TCL>.
15. FAOSTAT. FAOSTAT: Statistical database of Crop and livestock products for import and export quantity. FAO, Rome, Italy. Retrieved: 6 Sept 2018. (2018). Available at: <https://www.fao.org/faostat/en/#data/TCL>.
16. D'Amour, C. B., Wenz, L., Kalkuhl, M., Christoph Steckel, J. & Creutzig, F. Teleconnected food supply shocks. *Environ. Res. Lett.* **11**, 035007 (2016).
17. Brown, M. E. et al. Do markets and trade help or hurt the global food system adapt to climate change? *Food Policy* **68**, 154–159 (2017).
18. Distefano, T., Tuninetti, M., Laio, F. & Ridolfi, L. Tools for reconstructing the bilateral trade network: a critical assessment. *Econ. Syst. Res.* **32**, 378–394 (2019).
19. Tuninetti, M., Tamea, S., Laio, F. & Ridolfi, L. To trade or not to trade: Link prediction in the virtual water network. *Adv. Water Resour.* **110**, 528–537 (2017).
20. Fagiolo, G., Reyes, J. & Schiavo, S. On the topological properties of the world trade web: A weighted network analysis. *Phys. A Stat. Mech. its Appl.* **387**, 3868–3873 (2008).
21. Fagiolo, G., Reyes, J. & Schiavo, S. The evolution of the world trade web: A weighted-network analysis. *J. Evol. Econ.* **20**, 479–514 (2010).
22. Herman, P. R. Modeling complex network patterns in international trade. *Rev. World Econ.* **158**, 127–179 (2022).
23. Barua, S. & Valenzuela, E. Climate Change Impacts on Global Agricultural Trade Patterns: Evidence from the Past 50 Years. in Sixth International Conference on Sustainable Development 26–28 Sept 2018 (2018).
24. Dallmann, I. Weather Variations and International Trade. *Environ. Resour. Econ.* **72**, 155–206 (2019).
25. Mohan, P. Impact of Hurricanes on Agriculture: Evidence from the Caribbean. *Nat. Hazards Rev.* **18**, 04016012 (2017).
26. Osberghaus, D. The Effects of Natural Disasters and Weather Variations on International Trade and Financial Flows: a Review of the Empirical Literature. *Econ. Disasters Clim. Chang.* **3**, 305–325 (2019).
27. EMDAT. The OFDA/CRED International Disaster Database, University of Louvain (Accessed: 3 March 2022). (2022). Available at: <https://www.emdat.bel/>.
28. Lyubchich, V., Newlands, N. K., Ghahari, A., Mahdi, T. & Gel, Y. R. Insurance risk assessment in the face of climate change: Integrating data science and statistics. *WIREs Comput. Stat.* **11**, e1462 (2019).
29. Mills, E., Roth, R. & Lecomte, E. Availability and Affordability of Insurance Under Climate Change. A Growing Challenge for the U.S. *J. Insur. Regul.* **25**, (2005).
30. Willenbockel, D. & Oxfam. Extreme Weather Events and Crop Price Spikes in a Changing Climate. (2012).
31. Verma, M., Hertel, T. & Duffenbaugh, N. Market-oriented ethanol and corn-trade policies can reduce climate-induced US corn price volatility. *Environ. Res. Lett.* **9**, 064028 (2014).
32. Tigchelaar, M., Battisti, D. S., Naylor, R. L. & Ray, D. K. Future warming increases probability of globally synchronized maize production shocks. *Proc. Natl. Acad. Sci. USA.* **115**, 6644–6649 (2018).
33. Gaupp, F., Hall, J., Hochrainer-Stigler, S. & Dadson, S. Changing risks of simultaneous global breadbasket failure. *Nat. Clim. Chang.* **10**, 54–57 (2020).
34. Anderson, J. E. A theoretical foundation for the gravity equation. *Am. Econ. Rev.* **69**, 106–116 (1979).
35. Sohn, J. E. *The impact of aid for trade on the cost and time to trade: The case of Latin America and the Caribbean*. (United Nations Publications, 2013).
36. Bostan, I., Toderascu (Sandu), C. & Firtescu, B.-N. Exchange Rate Effects on International Commercial Trade Competitiveness. *Journal of Risk and Financial Management* **11**, 1–11 (2018).
37. Mohan, P. & Strobl, E. The economic impact of hurricanes in history: Evidence from sugar exports in the Caribbean from 1700 to 1960. *Weather. Clim. Soc.* **5**, 5–13 (2013).
38. Heger, M., Julca, A. & Paddison, O. Analysing the Impact of Natural Hazards in Small Economies The Caribbean Case. *Nat. Hazards* (2008).
39. De Benedictis, L. & Tajoli, L. The World Trade Network. *World Econ* **34**, 1417–1454 (2011).
40. Fagiolo, G. The international-trade network: Gravity equations and topological properties. *J. Econ. Interact. Coord* **5**, 1–25 (2010).
41. Herman, P. R. Essays on network effects in international trade and markets for differentiated goods. (2016).
42. Gutiérrez-Moya, E., Lozano, S. & Adenso-Díaz, B. Analysing the structure of the global wheat trade network: An ERGM approach. *Agronomy* **10**, 1–19 (2020).
43. Gopinath, M., Batarseh, F. A., Beckman, J., Kulkarni, A. & Jeong, S. International agricultural trade forecasting using machine learning. *Data Policy* **3**, 1–20 (2021).
44. Anderson, J. E. & Van Wincoop, E. Gravity with gravitas: A solution to the border puzzle. *Am. Econ. Rev.* **93**, 170–192 (2003).

45. Yotov, Y., Piermartini, R., Monteiro, J.-A. & Larch, M. *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. (2016). <https://doi.org/10.30875/abc0167e-en>
46. Mehrabi, Z. & Ramankutty, N. Synchronized failure of global crop production. *Nat. Ecol. Evol.* **3**, 780–786 (2019).
47. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl Acad. Sci.* **108**, 20260–20264 (2011).
48. Headey, D. Rethinking the global food crisis: The role of trade shocks. *Food Policy* **36**, 136–146 (2011).
49. Chen, C., Liaw, A. & Breiman, L. Using Random Forest to Learn Imbalanced Data. Available at: <https://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf>. University of California, Berkeley (2004).
50. Ke, X. Learning with Imbalanced Data. (2018).
51. Lee, S. K. & Jin, S. Decision tree approaches for zero-inflated count data. *J. Appl. Stat.* **33**, 853–865 (2006).
52. Mathlouthi, W., Larocque, D. & Fredette, M. Random forests for homogeneous and non-homogeneous Poisson processes with excess zeros. *Stat. Methods Med. Res.* **29**, 2217–2237 (2020).
53. Garcia-Marti, I., Zurita-Milla, R. & Swart, A. Modelling tick bite risk by combining random forests and count data regression models. *PLoS One* **14**, 1–22 (2019).
54. ECMWF. ERA5 hourly data on single levels from 1979 to present. *Reanalysis datasets* (2019).
55. Zhang, X. et al. Managing nitrogen for sustainable development. *Nature* **528**, 51–59 (2015).
56. Sacks, W. J., Deryng, D., Foley, J. A. & Ramankutty, N. Crop planting dates: An analysis of global patterns. *Glob. Ecol. Biogeogr.* **19**, 607–620 (2010).
57. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochem. Cycles* **22**, 1–2 (2008).
58. Mayer, T. & Zignago, S. Notes on CEPII's Distances Measures: The GeoDist Database. *SSRN Electron. J.* (2012). <https://doi.org/10.2139/ssrn.1994531>
59. Melitz, J. & Toubal, F. Native language, spoken language, translation and trade. *J. Int. Econ.* **93**, 351–363 (2014).
60. World Bank. World Development Indicators. (2018).
61. Head, K. & Mayer, T. Chapter 3 - Gravity Equations: Workhorse, Toolkit, and Cookbook. in *Handbook of International Economics* (eds. Gopinath, G., Helpman, E. & Rogoff, K. B. T.-H. of I. E.) **4**, 131–195 (Elsevier, 2014).
62. Head, K., Mayer, T. & Ries, J. The erosion of colonial trade linkages after independence. *J. Int. Econ.* **81**, 1–14 (2010).
63. Zhu, W., Porth, L. & Tan, K. S. A Credibility-based Yield Forecasting Model for Crop Reinsurance Pricing and Weather Risk Management. *Ssrn* **1**, 1–36 (2015).
64. Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M. & Morris, M. ergm: A package to fit, simulate and diagnose exponential-family models for networks. *J. Stat. Softw.* **24**, (2008).
65. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
66. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning. Elements of Statistical Learning* (2009). <https://doi.org/10.1007/978-0-387-84858-7>
67. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning*. (The MIT Press, 2016).
68. Beillouin, D., Schauburger, B., Bastos, A., Ciais, P. & Makowski, D. Impact of extreme weather conditions on European crop production in 2018. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190510 (2020).
69. Hoffman, A. L., Kemanian, A. R. & Forest, C. E. Analysis of climate signals in the crop yield record of sub-Saharan Africa. *Glob. Chang. Biol.* **24**, 143–157 (2018).
70. Jeong, J. H. et al. Random Forests for Global and Regional Crop Yield Predictions. *PLoS One* **11**, e0156571 (2016).
71. Altmann, A., Toloşi, L., Sander, O. & Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **26**, 1340–1347 (2010).

Acknowledgments

The funding support from the National Science Foundation CNS-1739823 is gratefully acknowledged.

Author contributions

S.V., V.L., and X.Z. designed research; S.V., V.L., and X.Z. analyzed data; and S.V., X.Z., and V.L. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43247-022-00591-7>.

Correspondence and requests for materials should be addressed to Xin Zhang.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Clare Davis.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022