

Preseason maize and wheat yield forecasts for early warning of crop failure

Received: 28 November 2023

Accepted: 9 August 2024

Published online: 23 August 2024



Weston Anderson^{1,2} , Shraddhanand Shukla³, Jim Verdin⁴, Andrew Hoell⁵, Christina Justice¹, Brian Barker¹, Kimberly Slinski^{1,2}, Nathan Lenssen^{6,7}, Jiale Lou⁸, Benjamin I. Cook^{9,10} & Amy McNally²

Provided the considerable logistical challenges of anticipatory action and disaster response programs, there is a need for early warning of crop failures at lead times of six to twelve months. But crop yield forecasts at these lead times are virtually nonexistent. By leveraging recent advances in climate forecasting, we demonstrate that global preseason crop yield forecasts are not only possible but are skillful over considerable portions of cropland. Globally, maize and wheat forecasts are skillful at lead times of up to a year ahead of harvest for 15% and 30% of harvested areas, respectively. Forecasts are most skillful in Southeast Africa and Southeast Asia for maize and parts of South and Central Asia, Australia, and Southeast South America for wheat. Wheat forecasts, furthermore, remain skillful at lead times of over 18 months ahead of harvest in some locations. Our results demonstrate that the potential for preseason crop yield forecasts is greater than previously appreciated.

Early warning systems that produce accurate and timely information provide vital decision support for effective humanitarian responses to food crises^{1,2}. Many of the programs implemented by government and non-governmental organizations (NGOs) to anticipate and respond to food crises require months to plan and execute. Examples include the distribution of farming inputs and drought-tolerant seed ahead of the growing season, securing the funds needed for disaster response, procurement and distribution of food aid, and scaling up of nutrition assistance programs^{3–6}. Anticipatory actions, such as the distribution of drought-tolerant seed varieties or switching to planting small grains on marginal lands instead of water-intensive crops can only be taken in advance of planting³. Even many disaster response policies, such as the shipment of in-kind food aid coming from the United States, to take one example, typically take around six months to reach target populations⁴. Likewise, the logistics of scaling up nutrition assistance programs can take months due to the time required to obtain the funds needed to finance such an operation, plan and conduct nutrition surveys to identify vulnerable populations, and set up food

distribution sites in targeted regions⁷. Even in the absence of exceptional funding needs, routine funding mechanisms for disaster response programs require months of advanced planning⁶.

Provided the considerable logistical challenges of anticipating and responding to disasters, there is a clear need for early warning information at lead times of six to twelve months. Crop yield forecasts made at these lead times could provide information on expected crop failures affecting the income of or food availability for vulnerable populations. But routine crop yield forecasts at lead times of a year do not currently exist. Research on crop yield forecast systems has been overwhelmingly focused on within-season forecasts, which are issued a few months prior to crop harvest, rather than preseason forecasts, which are issued prior to the start of the cropping season^{8,9}. The primary utility of preseason forecast systems is that they issue forecasts several months earlier than within-season forecast systems¹⁰. Preseason crop yield forecasts do exist and have been proposed for sugarcane in South Africa¹¹, maize in Zimbabwe¹², wheat and sugarcane in Australia^{13–15}, rice in the Philippines¹⁶, wheat, maize, and sugarbeet in

¹Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA. ²NASA Goddard Space Flight Center, Greenbelt, MD, USA.

³Climate Hazards Center, University of California-Santa Barbara, Santa Barbara, CA, USA. ⁴United States Agency for International Development, Washington, DC, USA. ⁵NOAA Physical Sciences Laboratory, Boulder, CO, USA. ⁶University of Colorado Boulder, Boulder, CO, USA. ⁷Colorado School of Mines, Golden, CO, USA. ⁸Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA. ⁹NASA Goddard Institute for Space Studies, New York, NY, USA. ¹⁰Lamont-Doherty Earth Observatory, Palisades, NY, USA. ✉ e-mail: weston@umd.edu

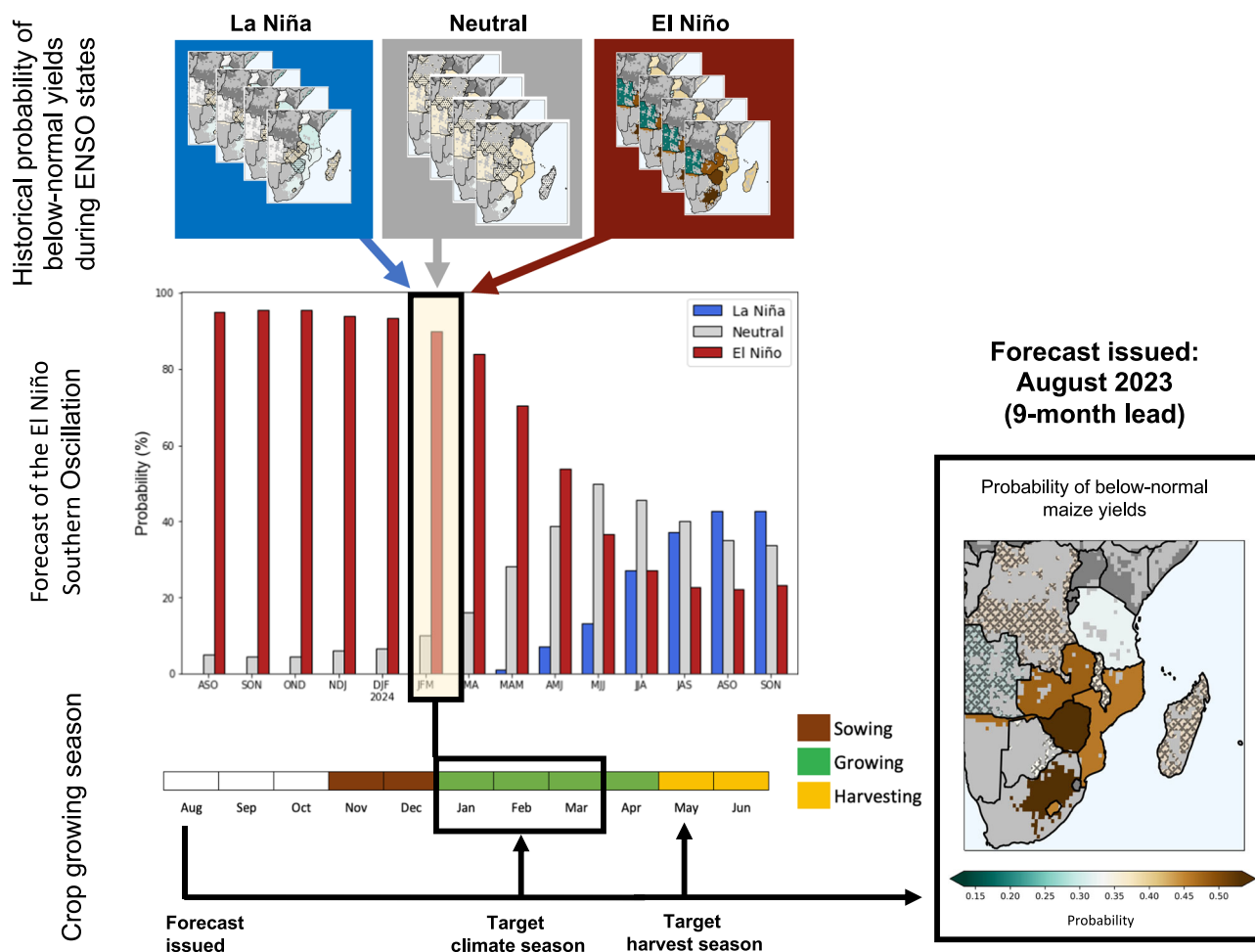


Fig. 1 | Schematic diagram of the preseason forecast model. For each target season during the growing season, historical observations of yield anomalies conditioned on the state of the ENSO in the target season are used to estimate the empirical probability of below-normal yields given the state of ENSO. These

probabilities are linearly combined using the probability of each ENSO state occurring during the target season to produce a single probability of below-normal yields given the ENSO forecast. Forecasts are only issued prior to the start of the vegetative season. Country boundaries are delineated using Natural Earth.

Europe^{17,18}, soybeans in the US¹⁹, and wheat, maize, rice, and soybeans at the global scale^{20–23}.

To date, however, even these preseason crop yield forecast systems generally focus on lead times of only six months prior to the start of harvest due to the availability of climate forecasts at these lead times^{21,22}. However recent advances in climate forecasts provide a means of extending preseason crop yield forecast lead times. In particular, skillful second-year forecasts of the El Niño Southern Oscillation (ENSO) have been developed using both machine-learning and model-analog-based methods^{24–26}. Provided the widespread influence of ENSO on crop yields globally^{27–30} and its use in past preseason crop yield forecast systems^{12,16,18,22,23}, this advance is highly relevant to the food security community.

In this article, we demonstrate that by leveraging the recent advances in ENSO forecasts, preseason crop yield forecasts are not only possible but are skillful over significant portions of cropland globally. In regions with strong ENSO teleconnections, both maize and wheat forecasts are skillful at lead times of up to a year ahead of harvest. Wheat forecasts, furthermore, remain skillful at lead times of over 18 months ahead of harvest in some countries. Our results demonstrate that the potential for preseason crop yield forecasts is greater than previously appreciated. Such forecasts require greater attention from the research community in order to support the next generation of operational food security early warning systems.

Results

Provided that the El Niño Southern Oscillation (ENSO) strongly affects maize and wheat yields in many parts of the world^{27,28,30,31} and that ENSO forecasts have multiyear skill^{24–26,32}, we propose a crop yield forecast system based on multiyear ENSO forecasts (Fig. 1). Our system combines the historical probability of below normal (bottom tercile) crop yield anomalies conditioned on ENSO phases with model-analog based ENSO forecasts³² during the crop growing season to forecast the likelihood that end-of-season yields will be below normal (see “Methods”).

Preseason forecast skill

Figures 2 and 3 demonstrate that crop yield forecasts based solely on ENSO forecasts can make skillful preseason forecasts out to and beyond a year prior to harvest for both maize and wheat. Forecasts are most skillful for maize in Southeast Africa and Southeast Asia while for wheat they are most skillful in parts of South and Central Asia, Australia, and Southeast South America. A slightly larger share of global wheat production is grown in countries for which skillful forecasts can be made as compared to maize (Fig. 3). For wheat forecasts at a lead of 6–9 months before harvest, countries having relative operating characteristics (ROC) scores of over 0.55, 0.6, and 0.65 represent 34%, 25%, and 18% of wheat harvested areas, respectively. For maize forecasts at a lead of 6–9 months before harvest, countries having ROC scores of over 0.55, 0.6, and 0.65 represent only 27%, 7%, and 2% of maize

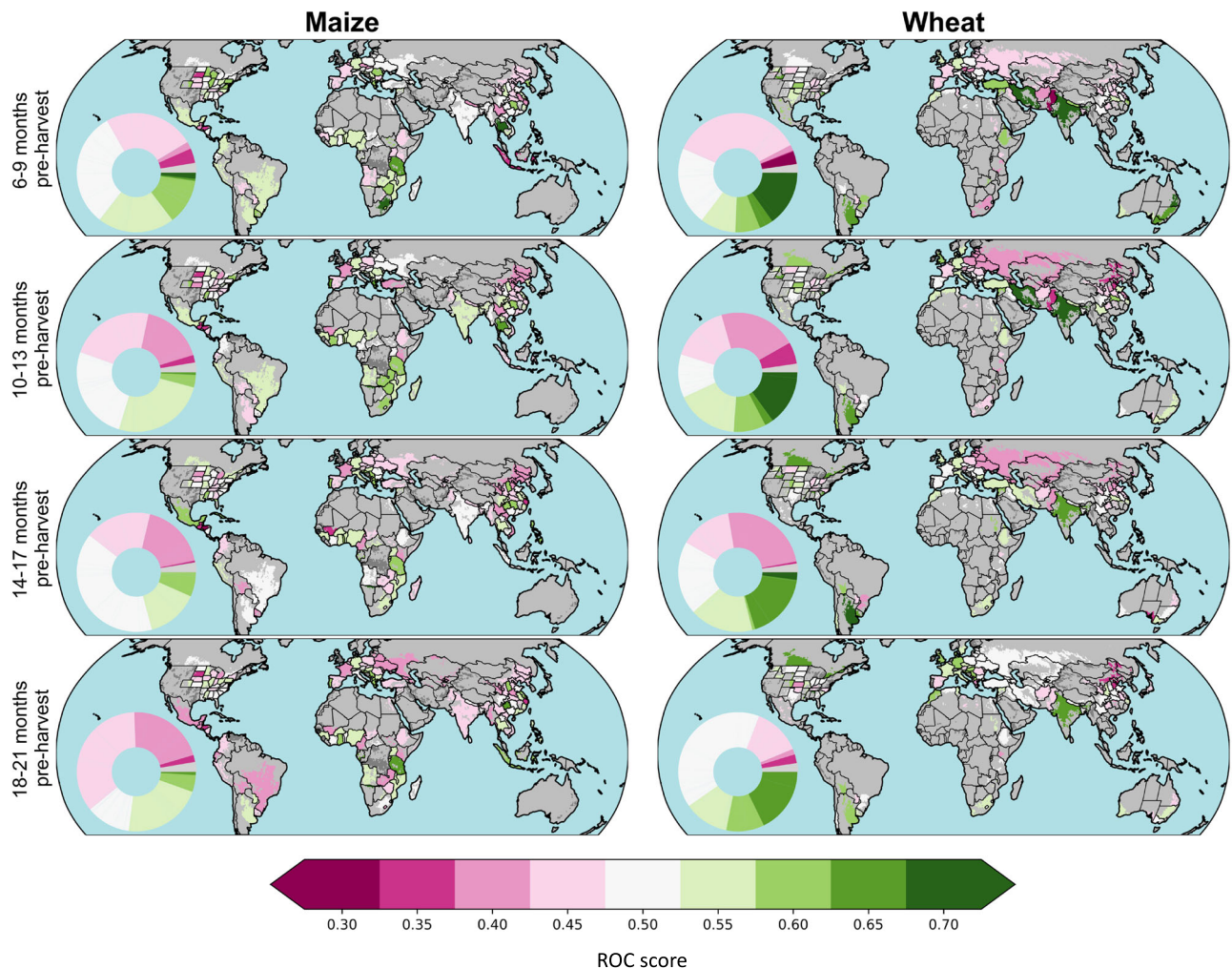


Fig. 2 | ROC score for lead times between six and twenty-one months ahead of harvest. ROC scores of one indicate a perfect model, while scores of less than 0.5 indicate a model with no skill. The pie chart insets in each figure indicate the ROC

scores in each country as percent of global production represented by that country. Country boundaries are delineated using Natural Earth, subnational boundaries are delineated using GADM.

harvested areas, respectively. Furthermore, skillful wheat forecasts ($\text{ROC} > 0.6$) tend to persist over ~20% of harvested areas out to leads of 18–21 months before harvest but forecasts with a $\text{ROC} > 0.65$ cover only ~5% of harvested areas for wheat at that lead time. Wheat and maize forecasts are skillful in roughly the same percentage of administrative units for which forecasts were made. Both wheat and maize forecasts have a $\text{ROC} > 0.6$ in ~8–12% of countries. That wheat forecasts are skillful over a greater fraction of harvested area is due to the concentration of wheat production in countries with forecast skill rather than the widespread presence of skill across wheat-producing countries. For wheat, the countries with $\text{ROC} > 0.6$ include major producing countries such as India, Australia, and Argentina, while for maize the most skillful forecasts are in Southeast Asia and Southeast Africa, which represent smaller portions of global maize production.

To better characterize forecast skills, we explore experiments that focus on the forecast lead time and the target season for which the forecast was made in individual countries. We find that the lead time at which forecast skill degrades depends on the country, crop, and month in which the forecast was issued (Fig. 4). Note that for each country the month in which the forecast is issued can also be described as the lead prior to harvest, so lead time can uniquely be converted into the calendar month in which the forecast was issued. For maize, forecast skill in South Africa and Thailand monotonically decline with increasing lead time until the forecasts become unskillful compared to

a random forecast ($\text{ROC} = 0.5$), which occurs at a lead time of around 20 months ahead of harvest (Fig. 4). In some wheat producing countries, on the other hand, forecast skill declines more slowly and non-monotonically.

Where forecasts are skillful, they are often reliable as well, meaning that outcomes are observed with the same frequency as the forecast probability. Figure 5 shows reliability diagrams for some of the countries with moderate to good skill for both maize and wheat at leads of 10–13 months, see SI Figs. 1 and 2 for the same figure at 6–9 and 18–21 months. For both maize and wheat, the frequency of the observed outcomes generally well matches the forecasted probability of that event in South Africa, Zimbabwe, India, Argentina, and Thailand. In Iran, however, forecasts tend to be under-confident. In addition to reliable forecasts, the forecasts are quite sharp—meaning that forecasts differ significantly from climatology. With the exception of Argentina, the system issues forecasts that range from below 20% to over 60% for the countries in Fig. 5. In Argentina the forecasts at 10–13 month leads are reliable but rarely differ much from the baseline probability of a poor year (33%).

To demonstrate that our conclusions are robust to the sampling period, we repeat our analysis in major maize and wheat breadbaskets using subnational data starting in 1931, the earliest date for which data was available in all breadbaskets (SI Note 2). SI Fig. 3 demonstrates that even in the 85-year record, our conclusions remain true: crop yield

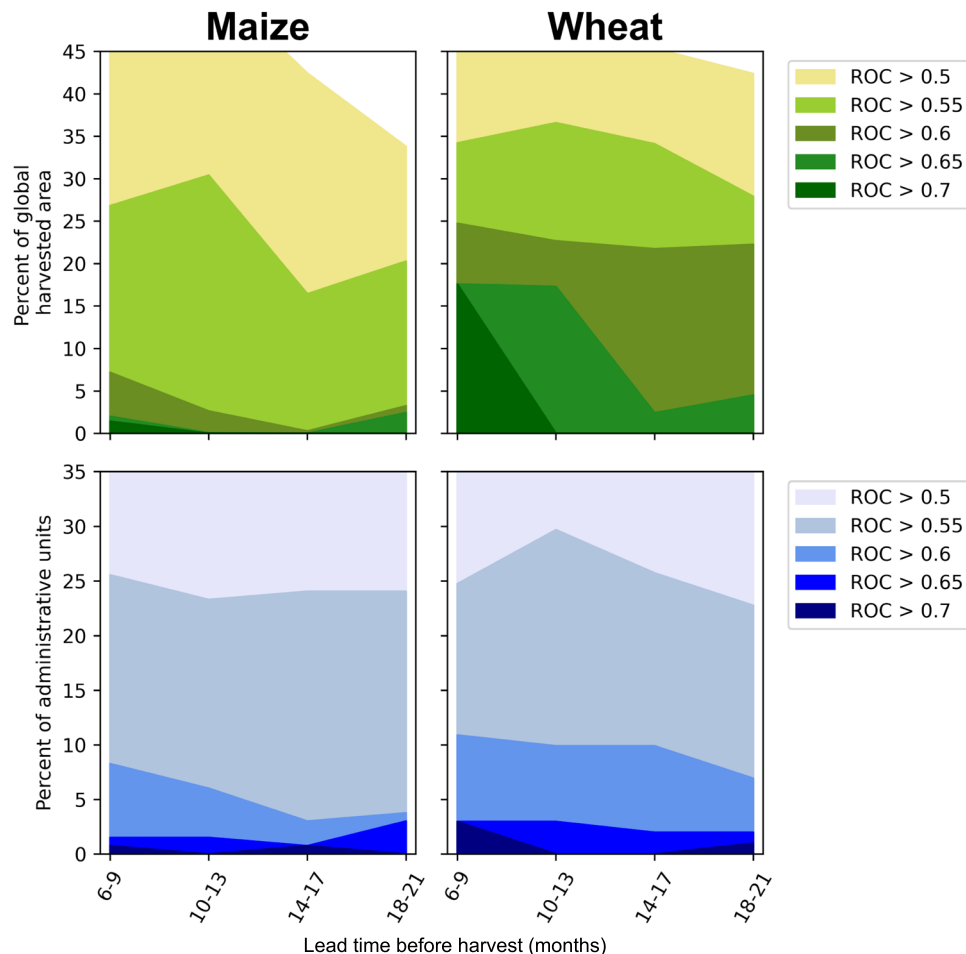


Fig. 3 | ROC scores for each lead time as a function of harvested area and number of countries. ROC scores of one indicate a perfect model, while scores of less than 0.5 indicate a model with no skill.

forecasts for both maize and wheat are skillful at lead times of 10–13 months ahead of harvest, and wheat forecasts are skillful at longer lead times than are maize forecasts.

Factors affecting forecast skill

Some of the variations in crop yield forecast skill are attributable to the model-analog-based ENSO forecasts, which demonstrate variations in skill both as a function of issue month and lead time. Model-analog-based ENSO forecasts issued immediately prior to the spring predictability barrier (e.g., those issued Feb–Apr) have lower skill than ENSO forecasts issued outside of this period³². We identify the crop yield forecasts based on ENSO forecasts issued immediately prior to the spring predictability barrier with gray shading in Fig. 4. As expected, these forecasts correspond to a period of rapid decline in the skill of crop yield forecasts.

The model-analog-based ENSO forecasts also have non-monotonic skill declines with increasing lead time³². The skill of ENSO forecasts issued in late summer, for example, declines rapidly from leads 6–9 before increasing again out to lead 15. This increase in skill at longer lead times may account for the increase in skill at longer leads in our crop yield forecasts. However, because the observed increase of skill at longer leads is not well understood in the ENSO forecasts themselves, further research on physical mechanisms of the re-emergence of forecast skill in both ENSO and crop yield forecasts is necessary.

Finally, crop yield forecast skill beyond a year may be related to enhanced year-two predictability of La Niña events that follow strong

El Niño events³³. The persistence of skill in India, in particular, may arise from the fact that La Niña events nearly always result in normal to above-normal yields (SI Fig. 4). This relationship is stronger in India than in other countries with notable ENSO teleconnections (e.g., maize in South Africa). The reliability diagrams for forecasts at lead times of 18–21 months (SI Fig. 3) further support this mechanism. The forecasts in India remain sharp and reliable but the sharpness skews towards low probabilities. The skillful and reliable long-lead prediction of a low probability of poor yields is, therefore, consistent with skillful year-two forecasts of La Niña events that preclude poor yields in India.

Because grain crop yields are more vulnerable to abiotic stress during particular portions of the growing season³⁴, it is possible that yield forecasts made using the climate conditions from those sensitive periods would be particularly skillful. To test this hypothesis, we evaluate the forecast skill separately for each 3-month climate season within the growing season rather than averaging them together as in previous analyses. Figure 4 confirms that forecast skill varies within the growing season for some countries. For example, in South Africa, forecasts for maize yields issued for the OND target season are consistently more skillful than those issued for later target seasons even at the same lead times. This target season-dependent skill is also apparent for maize in Thailand. For wheat, the dependence of forecast skill on target season is less clear although there is some relationship at longer lead times for India and Iran.

To test the extent to which our forecasts could be improved by better selecting target seasons or by improving ENSO forecasts themselves, we produce the following set of sensitivity experiments:

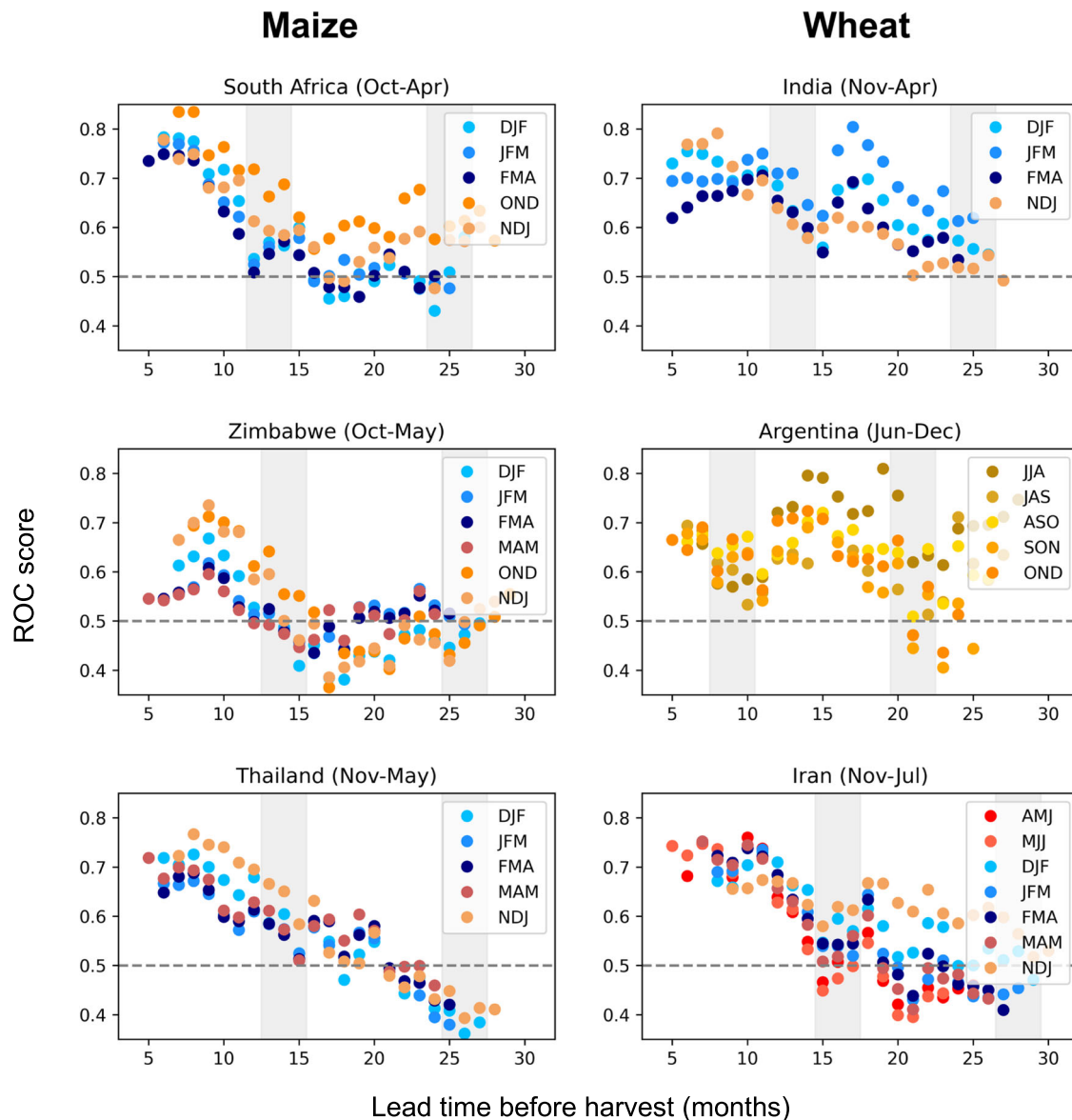


Fig. 4 | ROC scores for select countries at all lead times disaggregated by target season. ROC scores of one indicate a perfect model, while scores of less than 0.5 indicate a model with no skill.

(1) a ‘single season’ system in which a single target season is selected in a manner compatible with operational forecasts, (2) a ‘best season’ system in which the most skillful target season is selected a-posteriori, and (3) a ‘perfect ENSO forecast’ system that uses perfect deterministic information about ENSO phase rather than ENSO forecasts (see “Methods”). Figure 6 demonstrates that by selecting a single target season, even at lead times of 10–13 months before harvest, we can significantly increase the fraction of harvested areas with forecast skill ($AUC > 0.6$) in an operational context from 3% to 15% for maize and from 23% to 30% for wheat. While these improvements in skill are somewhat smaller than when the season that produces the highest skill is chosen a-posteriori, this is to be expected. And, notably, selecting a single season improves the forecast skill to comparable levels as being provided perfect information about the future categorical evolution of ENSO. By combining the a-posteriori season selection with perfect ENSO forecasts the system would provide skillful forecasts over 39% and 40% of maize and wheat harvested areas at a 10–13 month lead (Fig. 6). These numbers represent the ceiling for skillful forecasts using the present system.

Case study: a hindcast for the 1982 El Niño

Finally, to demonstrate the promise and limitations of an ENSO-based forecast system in real-time, we show hindcasts for maize during the 1982/83 El Niño event in Fig. 7. We choose this event because it has been identified as the largest synchronous maize failure in the recent historical period³⁵ and one in which ENSO played a major role organizing the global crop yield anomalies²⁷. At lead times of 10–13 months ahead of harvest, the forecast system predicts the greatest probability of below-normal yields in Southeast Africa and Southeast Asia, both of which ended up experiencing below-normal yields. At 14–17 months, however, the system does not capture the below-normal yields in West Africa or the United States. At this lead time, the system does not issue skillful forecasts that differ much from climatology despite the large crop yield anomalies that eventually develop throughout the Northern Hemisphere. This return to climatological forecast odds at the 14–17 month lead is related to the ENSO forecast itself (SI Fig. 5), which shows a ~40% chance of La Niña and ~30% chance of neutral or El Niño developing during that lead time. This example highlights two important points: (1) because many factors affect crop yields, even during the strongest ENSO events a forecast system based only on

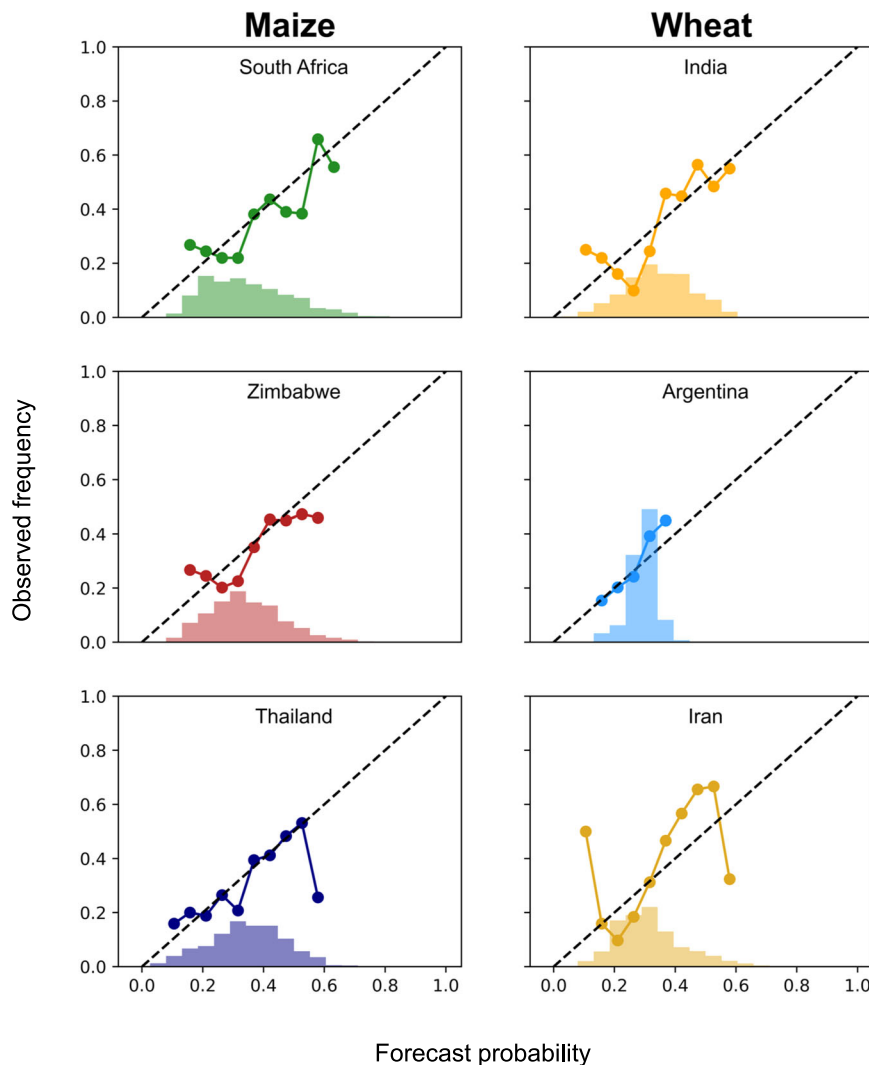


Fig. 5 | Reliability diagrams for select countries showing the observed frequency of below-normal yields vs the forecast probability of below-normal yields. The histograms in each panel show the frequency with which the forecast

system issues different probabilities. Reliability diagrams are shown for the 10–13 month lead for all panels.

ENSO will miss major crop yield failures and (2) an ENSO-based forecast system can be limited by the skill and sharpness of ENSO forecasts themselves in year two.

Discussion

This study finds that by leveraging recent advances in multiyear ENSO prediction, it's possible to make skillful forecasts of both maize and wheat yields at lead times up to a year ahead of harvest. Crop yield forecasts for maize and wheat issued 10–13 months before harvest had skill over 15% and 30% of global croplands, respectively (Fig. 6). Forecasts for wheat were more skillful over a larger fraction of global cropland at all lead times than those for maize (Figs. 2 and 3). The spatial pattern of forecast skill follows from the locations where ENSO most strongly affects crop yield variance²⁷, which is related to the strength, location, and timing of ENSO teleconnections relative to the climate-sensitive portions of the growing season^{29,30,36}. Forecasts were most skillful for maize in Southeast Africa and Southeast Asia while for wheat they are most skillful in India, parts of Asia, Australia, and Southeast South America.

We found that skillful forecasts could be made over a much greater proportion of the global harvested area if the climate conditions during specific portions of the growing season were used rather

than using the entire growing season (Figs. 4 and 6). We, therefore, recommend that future long-lead crop yield forecasting efforts consider identifying the portion of the growing season most relevant to connecting crop yields to climate conditions as a means of improving forecast skills.

Our study demonstrates both the promise and the limits of ENSO-based preseason crop yield forecasts. Even if provided with perfect information about the future evolution of ENSO, skillful forecasts cannot be made for more than ~40% of global harvested areas for both wheat and maize (Fig. 6). For wheat, the forecast skill of the current system is already nearing this limit at lead times of 10–13 months. From this we can conclude that improving year-one ENSO forecast skill is unlikely to greatly improve wheat yield forecasts, although as the case study demonstrates (Fig. 7 and Supplementary Information Fig. 5), improving year-two ENSO forecasts would improve long-lead crop yield forecasts. Based on these results, we recommend that preseason crop yield forecast systems incorporate modes of variability beyond ENSO to expand the area over which skillful forecasts can be made at lead times out to a year, while also recognizing that the ongoing improvements in year-two ENSO forecast skill will directly translate into improved year-two crop yield forecasts.

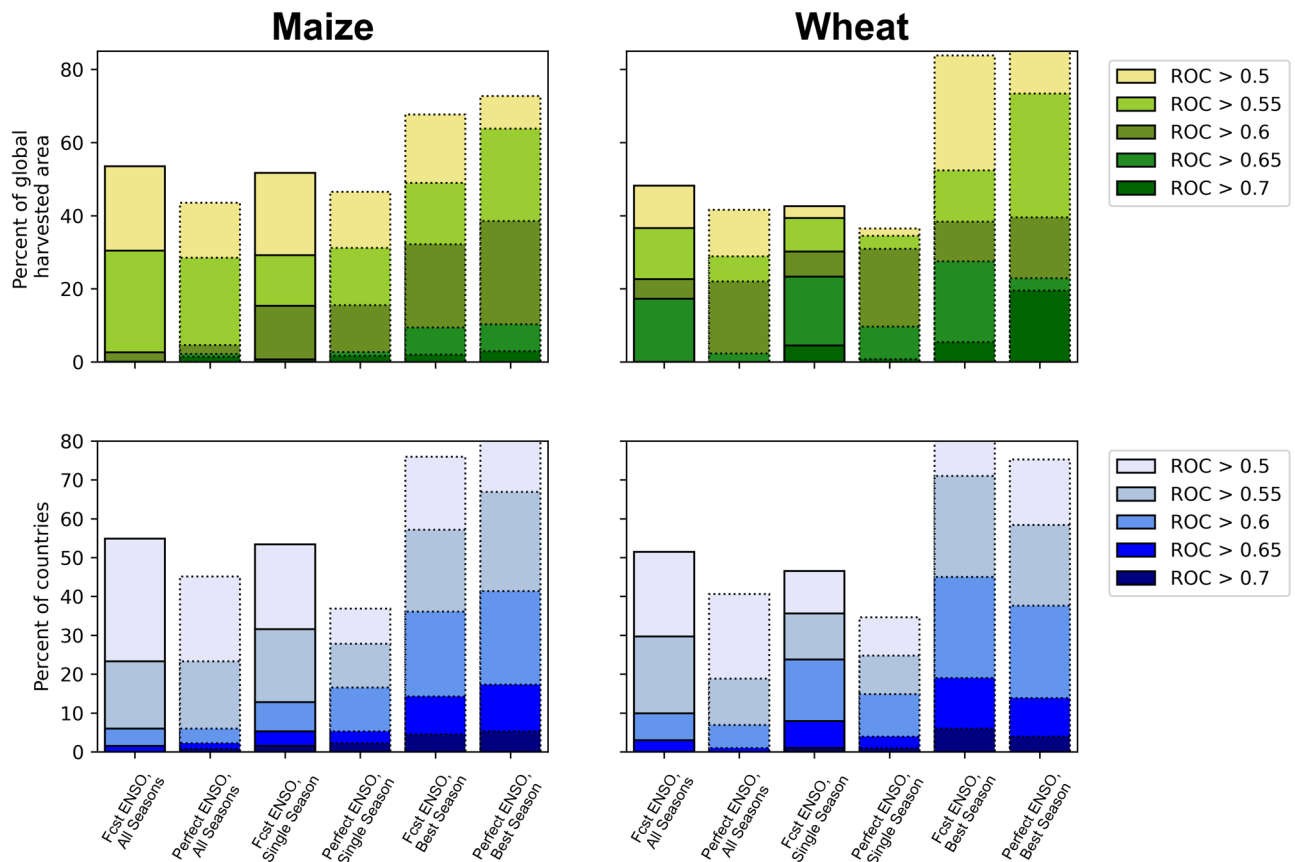


Fig. 6 | The ROC score at leads of 10–13 months before harvest. Each bar shows a combination of forecast systems that select all target seasons (all seasons), select an optimal target season in a manner consistent with operational forecasting (single season), or select the most skillful target season a-posteriori (best season) with

either forecasts of ENSO (Fcst ENSO) or perfect deterministic information about the future evolution of ENSO (perfect ENSO). Solid lines denote different operational model configurations while dotted lines denote model configurations used to explore the theoretical limits of the forecast system.

Whether preseason crop yield forecasts are presently skillful enough to inform decisions will depend on the region and lead time of the forecast, the proposed action, and the risk tolerance of the decision maker. The relatively higher forecast skill in Southern Africa, Australia, and parts of Asia may be suitable for anticipatory or response-related actions. But in regions where preseason forecasts have only marginal skill, the system is likely better suited to the ready-set-go framework for using seasonal forecasts³⁷. This approach identifies low-cost actions—such as the updating of contingency plans (ready)—that can be taken using long-lead outlooks of a potential hazard. As the anticipated impact draws nearer and more skillful forecasts become available, local agencies may then prepare to implement those contingency plans (set) so that once the hazard arrives, appropriate action can be taken in a timely manner (go).

Finally, our analysis has a number of limitations that should be considered when interpreting the results. We build our crop yield forecast system mostly on country-level Food and Agriculture Organization (FAO) crop yield statistics, which have known issues such as the random errors introduced by a reduced capacity of the FAO to accurately collect country-level data in Africa³⁸ and the systematic errors introduced by over-reporting of achievements³⁹. We minimize these effects through our quality control procedures, by investigating only deviations from long-term trends, and by using as many ENSO years as possible in formulating and evaluating our model. However, our results should still be considered in light of the limitations of imperfect data.

The research community has not yet identified the intrinsic limits of skillful preseason crop yield forecasts. We demonstrate this by using an ENSO-based forecast system to produce skillful crop yield forecasts

at lead times of over a year ahead of harvest. That a forecast system as simple as ours produces skillful forecasts at unprecedented lead times illustrates the extent to which preseason crop yield forecasts have been neglected in the research community. Provided recent advances in both climate forecasting methods and crop yield forecasting methods, there is significant potential for developing preseason crop yield forecast systems that are skillful enough to be a useful component of early warning systems.

Methods

Crop yield statistics

We use crop yield data from both the 20th Century Crop Yield Statistics v1 (see SI Note 2), which consists of subnational data from 1920–2017 in all major wheat and maize bread baskets for which data was available^{31,40}, and the FAO country-level crop yield data from 1961 to 2021⁴¹. We chose these two datasets because they are available for relatively long time periods, which is critical to training and evaluating the performance of our model.

Crop calendars and crop masks

To identify the location of crop growing areas we use the GEOGLAM crop masks for maize and wheat⁴². For crop calendars, we use the vegetative to the reproductive stage of the GEOGLAM main season for both maize and wheat. The exception to this is Brazil, Kenya, and Somalia. In Brazil, we use the second maize season because it was, for the majority of our study period, the primary maize production season until the second safrinha-cycle planting became the larger of the two production seasons around 2010²⁹. In Kenya and Somalia, we use the short-rains season, which is the smaller of the two maize production

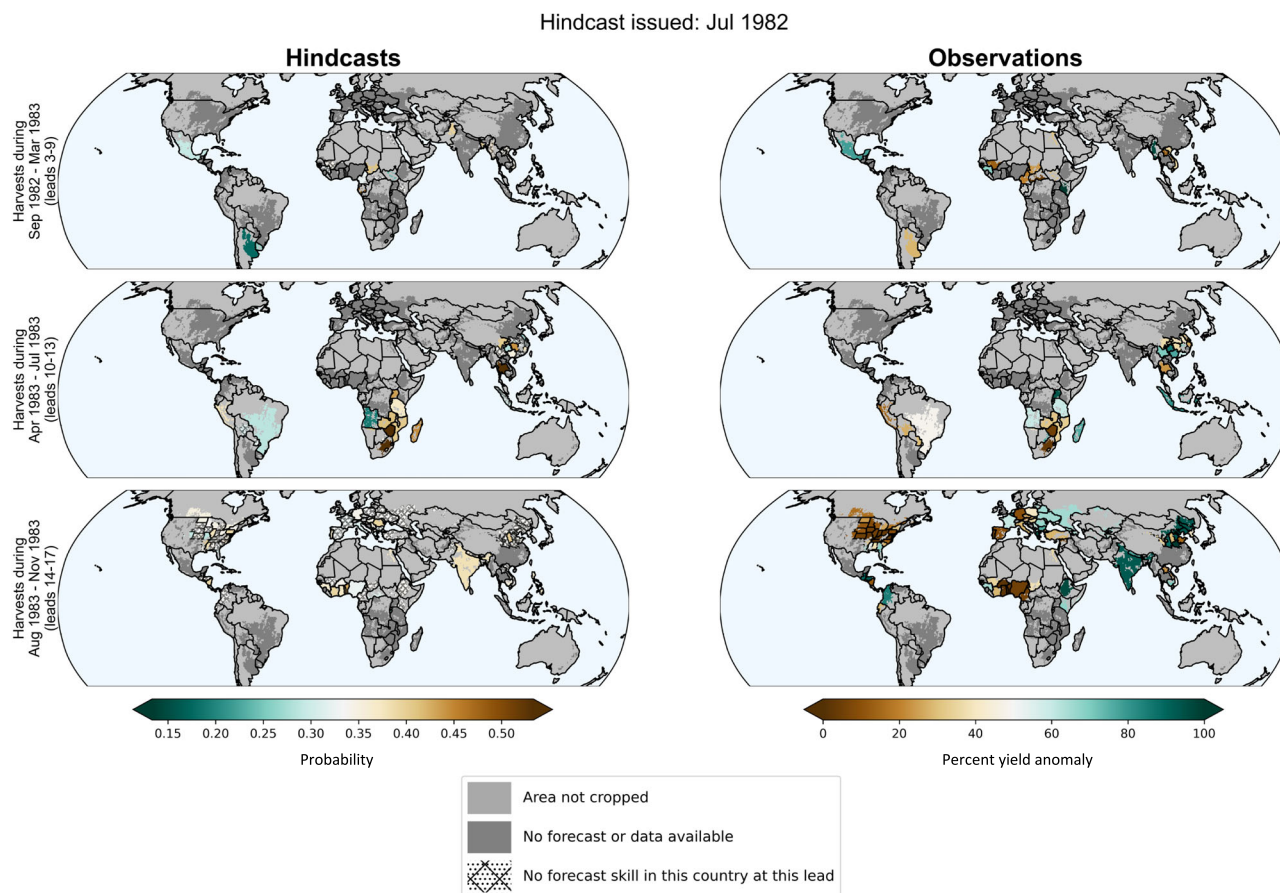


Fig. 7 | Hindcast for maize of the probability of below normal yields issued July 1982 (left panels) and observations of percent yield anomalies that eventually transpired (right panels). Light gray indicates that an area is not cropped, and dark gray indicates that an area has no forecast issued for the given forecast dates

or that historical data was not available for the period in the forecast range. Hatched areas in the hindcasts indicate no skill for that forecast lead and country (ROC score < 0.5). Country boundaries are delineated using Natural Earth, subnational boundaries are delineated using GADM.

seasons, as the target season because it is strongly affected by ENSO and generally has greater interannual variability as compared to the long rains season. For wheat, we use the winter wheat season from GEOGLAM as the main season in all countries except Canada and Yemen, for which we use the spring wheat growing season as the main season.

Observations and forecasts of ENSO

To identify the historical evolution of ENSO we use the Oceanic Niño Index (ONI), which measures anomalies in the sea surface temperature (SST) in the Niño 3.4 region (5°N–5°S, 120°–170°W) in the ERSSTv5 dataset⁴³. We calculate anomalies using a 30-year “fair-sliding” climatology behind the forecast period so that model performance is evaluated in a manner consistent with a routine forecast system⁴⁴ as in ref. 32. For ENSO forecasts we use the NOAA/PSL and University of Colorado/CIRES “model-analog” forecasts, which have comparable skill to operational initialized forecast models^{24,25}. The benefits of a model-analog approach include the availability of an extensive hindcast record dating back to the 1850s and the availability of forecasts at leads up to two years³². The long time series and extensive set of hindcasts allow us to evaluate forecasts for all years for which crop yield data is available.

Crop yield anomaly time series

Because the FAO dataset provides estimates of crop yields even when data is not available, some quality screening of the data is necessary as a first step. Any yield value in the FAO dataset with the

data quality flag “imputed value” or “missing value” is removed from our analysis. We then screen the FAO data to remove values with unrealistically low variance relative to a trend based on second differences of the yield time series. This was necessary because at times FAO data appears to be filled using a trend (e.g., a constant second-difference) without an associated data quality flag (see SI Note 1 and SI Fig. 6). We flag any of the second differences in which four consecutive values were less than 50 kg/ha as being unrealistically low. This is equivalent to identifying at least five values that didn’t meaningfully deviate from the yield trend. While the exact threshold used is subjective, we found 50 kg/ha to be a threshold that removed many of the most likely in-filled values without erroneously removing values in low-variance production systems with higher data quality, such as heavily irrigated wheat in India. SI Fig. 6 shows six examples of countries with values that were flagged and removed (highlighted areas) for wheat and maize. Finally, we correct a reporting issue for wheat in Argentina (see Supplementary Information). Following quality controls, we remove any country that has fewer than five years of data remaining.

To produce continuous, homogenous time series it is necessary to combine a number of countries from the FAO data into aggregate groups to account for changes in national boundaries over time. In particular, we group the following nations together: the countries that make up the former USSR, Sudan and South Sudan, Yugoslavia, Belgium-Luxembourg, and Czechoslovakia. Following this procedure, we recompute aggregate yields by dividing reported production by reported harvested area.

We supplement the FAO data with subnational data in the United States, China, and Australia with data collated in the 20th Century Crop Yield Statistics v1^{31,40} and updated to 2020. We use total wheat statistics in China and the United States to match the FAO data, which presents data only for combined seasons and combined sub-crops.

For both crop yield datasets, we use a low-frequency Gaussian filter with a kernel standard deviation of three years to calculate expected yields. We then calculate the percent yield anomalies as below:

$$Ya_i = 100 * (Y_i - Ye_i) / (Ye_i) \quad (1)$$

where Ya_i is the yield anomaly in year i , Y_i is the yield value in year i , and Ye_i is the expected yield in the year i . We find that our results remain largely the same when using a five-year running mean to define the expected yield instead of a low-frequency Gaussian filter (results not shown). We use percent yield anomalies rather than absolute yield anomalies to account for the increase in yield variance that accompanies increases in absolute yields over time⁴⁵. To align the crop harvest season with the ENSO forecasts, we identify the growing seasons for each crop in each country that span the calendar year to ensure that forecasts made for those seasons at the end of one calendar year correctly align with the yield value, which is reported in FAO statistics as the harvest year.

Climate forecasts

Our climate forecasts are only issued prior to the start of the growing season, meaning we do not produce forecasts for crops and countries in which the vegetative season has already begun. We refer to our forecasts as 'preseason' to differentiate them from crop yield forecasts made for end-of-season yields based on observations of moisture availability and vegetative health during the growing season. We choose to exclusively produce preseason forecasts rather than continuing to issue forecasts throughout the growing season because we anticipate that the usefulness of climate forecast-based yield predictions will be primarily prior to the start of the vegetative phase and/or the rainy season. Once the growing season is significantly underway, we expect earth observation-based forecasts to be a more accurate representation of the likely end-of-season yield outcome in a country. A consequence of this decision is that nearly all of our forecasts are at leads longer than five months prior to harvest. This distinguishes our forecasts from past efforts, which make forecasts for a crop up until harvest.

For a given ENSO forecast in a given target season, we first identify the countries that will have crops that are in-season during the target season, which we define as countries with the majority (>50% of the harvested area) of the main season maize or wheat in the vegetative to reproductive phase using the GEOGLAM crop calendars and crop masks. Note that we exclude countries where the crop is currently in season even if it will also be in season during the target season. We choose the vegetative to reproductive period because grain crops are particularly vulnerable to abiotic stresses during the reproductive and grain-filling stages of crop growth³⁴.

For each country with a valid in-season crop, we produce a forecast of the probability of below-normal crop yields as follows: (1) separate historical crop yield anomalies into terciles corresponding to above-normal, normal, and below normal, (2) identify historical years in which ENSO was active in each phase (El Niño, La Niña, neutral) for the given target season, (3) separate the country's crop yield anomalies according to the ENSO phase in each year of the historical record to calculate the probability that the crop yield is in the below normal tercile given the ENSO phase, and (4) linearly combine each of the three conditional probabilities proportional to the forecast probability of that phase of ENSO occurring in the target season. This can be

written as follows:

$$Pybn_i = (pEN_i) * (Pybn|EN) + (pN_i) * (Pybn|N) + (pLN_i) * (Pybn|LN) \quad (2)$$

where $Pybn_i$ is, for target season i , the issued forecast probability that the yield is in the below normal tercile, pEN_i is the probability that ENSO is in the El Niño phase during target season i , and $(Pybn|EN)$ is the historical probability that yields are below normal given that ENSO is in the El Niño phase during the 3-month target season i . For each country and crop in a given year, the system issues forecasts for all overlapping 3-month target seasons within the vegetative period of the growing season. This is done to reflect the forecasted evolution of the tropical Pacific throughout the vegetative season. For each lead prior to harvest, therefore, we average all forecasts issued for the growing season so that the system issues a single forecast probability for a growing season.

Note that throughout this article we refer to lead time as the time between when the forecast is issued and the start of harvest, not the time between when a forecast is issued and the target season (see Fig. 1). For example, two forecasts issued in August with target seasons of JFM and FMA for a harvest occurring in May will both have a lead time of nine months. We use this notation because the target quantity of our forecasts is the end-of-season crop yields, not the mid-growing-season ENSO state on which the forecasts are conditioned.

Conceptually, our model can be considered as a formalization of the expectation for below-normal crop yields based on a linear combination of historical observations relative to the forecasted evolution of ENSO. While conceptually straightforward, it provides a powerful benchmark forecast to take advantage of the known predictability of ENSO and of known ENSO teleconnections to crop yields while also accounting for the uncertainty of the evolution of the tropical Pacific on multiyear time scales.

Skill evaluation

We use a probabilistic categorical model for our preseason forecasts because such a model recognizes that although forecasts of precise crop yields are likely impossible at such long leads, there may still be predictable shifts in the distribution of yield outcomes. Such probabilistic categorical models are widespread in the climate forecasting community, but are not commonly applied to crop yield forecasts. A further advantage of such categorical forecasts is that they account for the low precision of FAO crop yield estimates in many parts of the world by requiring only that the observations are precise enough to place the year in the correct tercile.

We evaluate forecast skill in a hold-one-out cross-validation approach using the ROC score, which measures the ratio of the hit rate to the corresponding false alarm rate of a categorical forecast⁴⁶. The area under the ROC curve is a measure of forecast discrimination or, put another way, a measure of whether a forecast differs given different outcomes. A ROC score of 0.5 or lower indicates a forecast model with no skill, while a score of 1 indicates a perfect forecast model. We follow the conventions of ref. 21, who previously evaluated probabilistic global crop yield forecasts, in defining forecasts with an ROC > 0.6 as having skill. For countries with skillful forecasts, we furthermore present metrics on forecast reliability, which measures whether observed outcomes are conditioned on issued forecast probabilities. A reliability diagram plots the forecast probability against the frequency of an observed outcome. All analyses of reliability and skill evaluations are conducted separately for wheat and maize in each country at a variety of lead times.

Single-season and best-season experiments

To test the improvement in skill when targeting a specific season within the growing season we configure the model to identify and use ENSO forecasts for a single target season within the growing season

rather than using all target seasons. We do so within the leave-one-out cross-validation framework by using only the training data (not the withheld test year) to correlate observed crop yields with observed Niño 3.4 SST anomalies using a Spearman rank correlation for each three-month season in the growing season. We then use the season with the greatest absolute correlation that was determined from our training data as our single target season to predict the test data year.

To test the potential improvements to be gained by selecting a single target season rather than using all seasons we conduct an a-posteriori evaluation of the all-seasons model by selecting a single target season for each country based on the ROC score of the model using each individual target season at leads of 6–13 months. This a-posteriori analysis represents the maximum potential benefit of using a single target season configuration in the model system.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data needed to reproduce the analysis is publicly available. FAO-STAT is available from: <https://www.fao.org/faostat/en/>, the 20th-century crop yield data is available: <https://sedac.ciesin.columbia.edu/data/set/food-twentieth-century-crop-statistics-1900-2017>, the ENSO model analogs are available from: <https://psl.noaa.gov/enso/forecasts.html>. The crop masks used in the analysis (GEOGLAM BACs) are available from: <https://zenodo.org/records/7230863>.

Code availability

Code to analyze the data was written in Python 3.8.8 and is available from: <https://github.com/WestonAnderson/yld-fcsts>. All spatial plots were made using Cartopy⁴⁷.

References

- Funk, C. et al. Recognizing the famine early warning systems network: over 30 years of drought early warning science advances and partnerships promoting global food security. *Bull. Am. Meteorol. Soc.* **100**, 1011–1027 (2019).
- Funk, C. et al. Tailored forecasts can predict extreme climate informing proactive interventions in east africa. *Earth's Future* **11**, e2023EF003524 (2023).
- Mugiyo, H. et al. El niño's effects on southern african agriculture in 2023/24 and anticipatory action strategies to reduce the impacts in zimbabwe. *Atmosphere* **14**, 1692 (2023).
- US Government Accountability Office. *International food aid: pre-positioning speeds delivery of emergency aid, but additional monitoring of time frames and costs is needed. Report to the Chairwoman, Committee on Agriculture, Nutrition, and Forestry, US Senate* (US Government Accountability Office, 2014).
- Maxwell, D., Majid, N., Adan, G., Abdirahman, K. & Kim, J. J. Facing famine: somali experiences in the famine of 2011. *Food Policy* **65**, 63–73 (2016).
- Smith, R. *Ocha on message: consolidated appeal process*. Tech. Rep., UN Office for Coordination of Humanitarian Affairs (OCHA, 2012).
- Maxwell, N., Daniel G. Majid. *Famine in Somalia: Competing Imperatives Collective Failures, 2011–2012* (Oxford University Press, 2016).
- Schauberger, B., Jägermeyr, J. & Gornott, C. A systematic review of local to regional yield forecasting approaches and frequently used data resources. *Eur. J. Agron.* **120**, 126153 (2020).
- Basso, B. & Liu, L. Seasonal crop yield forecast: methods, applications, and accuracies. *Adv. Agron.* **154**, 201–255 (2019).
- Iizumi, T. et al. Evaluating the 2019 NARO-APCC joint crop forecasting service yield forecasts for northern hemisphere countries. *Weather Forecast.* **36**, 879–891 (2021).
- Bezuidenhout, C. & Schulze, R. Application of seasonal climate outlooks to forecast sugarcane production in south africa. *Clim. Res.* **30**, 239–246 (2006).
- Cane, M. A., Eshel, G. & Buckland, R. W. Forecasting zimbabwean maize yield using eastern equatorial pacific sea surface temperature. *Nature* **370**, 204–205 (1994).
- Yuan, C. & Yamagata, T. Impacts of IOD, ENSO and ENSO modoki on the australian winter wheat yields in recent decades. *Sci. Rep.* **5**, 17252 (2015).
- Mavromatis, T. Pre-season prediction of regional rainfed wheat yield in northern greece with ceres-wheat. *Theor. Appl. Climatol.* **117**, 653–665 (2014).
- Everingham, Y., Sexton, J., Skocaj, D. & Inman-Bamber, G. Accurate prediction of sugarcane yield using a random forest algorithm. *Agron. Sustain. Dev.* **36**, 1–9 (2016).
- Koide, N. et al. Prediction of rice production in the philippines using seasonal climate forecasts. *J. Appl. Meteorol. Climatol.* **52**, 552–569 (2013).
- Ceglar, A., Turco, M., Toreti, A. & Doblas-Reyes, F. J. Linking crop yield anomalies to large-scale atmospheric circulation in europe. *Agric. For. Meteorol.* **240**, 35–45 (2017).
- Guimarães Nobre, G., Hunink, J. E., Baruth, B., Aerts, J. C. & Ward, P. J. Translating large-scale climate variability into crop production forecast in europe. *Sci. Rep.* **9**, 1277 (2019).
- Vijverberg, S., Hamed, R. & Coumou, D. Skillful us soy yield forecasts at presowing lead times. *Artif. Intell. Earth Syst.* **2**, e210009 (2023).
- Iizumi, T., Takaya, Y., Kim, W., Nakaegawa, T. & Maeda, S. Global within-season yield anomaly prediction for major crops derived using seasonal forecasts of large-scale climate indices and regional temperature and precipitation. *Weather Forecast.* **36**, 285–299 (2021).
- Iizumi, T., Shin, Y., Kim, W., Kim, M. & Choi, J. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. *Clim. Serv.* **11**, 13–23 (2018).
- Iizumi, T. et al. Prediction of seasonal climate-induced variations in global food production. *Nat. Clim. Change* **3**, 904–908 (2013).
- Cao, J. et al. Forecasting global crop yields based on el nino southern oscillation early signals. *Agric. Syst.* **205**, 103564 (2023).
- Ding, H., Newman, M., Alexander, M. A. & Wittenberg, A. T. Skillful climate forecasts of the tropical indo-pacific ocean using model-analogs. *J. Clim.* **31**, 5437–5459 (2018).
- Ding, H., Newman, M., Alexander, M. A. & Wittenberg, A. T. Diagnosing secular variations in retrospective enso seasonal forecast skill using cmip5 model-analogs. *Geophys. Res. Lett.* **46**, 1721–1730 (2019).
- Ham, Y.-G., Kim, J.-H. & Luo, J.-J. Deep learning for multi-year enso forecasts. *Nature* **573**, 568–572 (2019).
- Anderson, W., Seager, R., Baethgen, W., Cane, M. & You, L. Synchronous crop failures and climate-forced production variability. *Sci. Adv.* **5**, eaaw1976 (2019).
- Iizumi, T. et al. Impacts of El Niño Southern Oscillation on the global yields of major crops. *Nat. Commun.* **5**, 3712 (2014).
- Anderson, W., Seager, R., Baethgen, W. & Cane, M. Crop production variability in north and south america forced by life-cycles of the el niño southern oscillation. *Agric. For. Meteorol.* **239**, 151–165 (2017).
- Anderson, W., Seager, R., Baethgen, W. & Cane, M. Trans-pacific enso teleconnections pose a correlated risk to agriculture. *Agric. For. Meteorol.* **262**, 298–309 (2018).
- Anderson, W. et al. Climate variability and simultaneous bread-basket yield shocks as observed in long-term yield records. *Agric. For. Meteorol.* **331**, 109321 (2023).

32. Lou, J., Newman, M. & Hoell, A. Multi-decadal variation of enso forecast skill since the late 1800s. *npj Clim. Atmos. Sci.* **6**, 89 (2023).
33. Lenssen, N. et al. Strong el niño events lead to robust multi-year ENSO predictability. *Geophys. Res. Lett.* **51**, e2023GL106988 (2024).
34. Prasad, P., Staggenborg, S. & Ristic, Z. Impacts of drought and/or heat stress on physiological, developmental, growth, and yield processes of crop plants. *Adv. Agric. Syst. Model. Series* **1**, 301–355 (2008).
35. Mehrabi, Z. & Ramankutty, N. Synchronized failure of global crop production. *Nat. Ecol. Evol.* **3**, 780–786 (2019).
36. Lenssen, N. J., Goddard, L. & Mason, S. Seasonal forecast skill of enso teleconnection maps. *Weather Forecast.* **35**, 2387–2406 (2020).
37. Goddard, L., Baethgen, W. E., Bhojwani, H. & Robertson, A. W. The international research institute for climate & society: why, what and how. *Earth Perspect.* **1**, 1–14 (2014).
38. Dunmore, J. & Karlsson, J. Independent evaluation of fao's role and work in statistics. *Food Agric. Organ. Rome* **77** <https://www.fao.org/4/k2921e/k2921e.pdf> (2008).
39. Xiao, K. & Womack, B. Distortion and credibility within china's internal information system. *J. Contemp. China* **23**, 680–697 (2014).
40. Anderson, W. et al. Twentieth century crop statistics, 1900–2017. NASA Socioeconomic Data and Applications Center 495 (SEDAC). *Tech. Rep.* <https://doi.org/10.7927/tmsp-sg82> (2022).
41. FAO. Tracking progress on food and agriculture-related SDG indicators 2022. <https://doi.org/10.4060/cc1403en> (Rome, 2022).
42. Becker-Reshef, I. et al. Crop type maps for operational global agricultural monitoring. *Sci. Data* **10**, 172 (2023).
43. Huang, B. et al. Extended reconstructed sea surface temperature, version 5 (ersstv5): upgrades, validations, and intercomparisons. *J. Clim.* **30**, 8179–8205 (2017).
44. Risbey, J. S. et al. Standard assessments of climate forecast skill can be misleading. *Nat. Commun.* **12**, 4346 (2021).
45. Calderini, D. F. & Slafer, G. A. Changes in yield and yield stability in wheat during the 20th century. *Field Crops Res.* **57**, 335–347 (1998).
46. Kharin, V. V. & Zwiers, F. W. On the roc score of probability forecasts. *J. Clim.* **16**, 4145–4150 (2003).
47. Met Office Cartopy: a cartographic python library with a Matplotlib interface. *Exeter, Devon* <https://scitools.org.uk/cartopy> (2010–2015).

Acknowledgements

A.M., K.S., W.A., and B.C. were supported by the Bureau of Humanitarian Assistance, U.S. Agency for International Development, under the terms of NASA GSFC PAPA AID-720BHAH00005 Famine Early Warning Systems Network (FEWS NET). S.S. was supported by the U.S. Geological

Survey, United States grant G21AC00026. J.L. was primarily supported by USAID's Famine Early Warning Systems Network, AID-OFDA-T-17-00002.

Author contributions

W.A., S.S., J.V., A.H., C.J., B.C., K.S., and B.B. designed the analysis. W.A., N.L., and J.L. ran the analysis. W.A., S.S., J.V., A.H., C.J., B.C., B.B., K.S., and A.M. helped to analyze the results and write the manuscript.

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/s41467-024-51555-8>.

Correspondence and requests for materials should be addressed to Weston Anderson.

Peer review information *Nature Communications* thanks Marijn van der Velde and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024