LETTER • OPEN ACCESS

Weather dataset choice introduces uncertainty to estimates of crop yield responses to climate variability and change

To cite this article: B Parkes *et al* 2019 *Environ. Res. Lett.* **14** 124089

View the article online for updates and enhancements.

Environmental Research Letters

LETTER

Weather dataset choice introduces uncertainty to estimates of crop yield responses to climate variability and change

B Parkes[1,4,5] , T P Higginbottom[1] , K Hufkens[3] , F Ceballos[2] , B Kramer[2] and T Foster[1,4]

[1] Department of Mechanical, Aerospace and Civil Engineering, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
[2] International Food Policy Research Institute, 1201 Eye St, NW, Washington, DC 20005-3915, United States of America
[3] Computational & Applied Vegetation Ecology Lab, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, Ghent, B-9000, Belgium
[4] Centre for Crisis Studies and Mitigation University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
[5] Author to whom any correspondence should be addressed.

E-mail: ben.parkes@manchester.ac.uk

Abstract

Weather shocks, such as heatwaves, droughts, and excess rainfall, are a major cause of crop yield losses and food insecurity worldwide. Statistical or process-based crop models can be used to quantify how yields will respond to these events and future climate change. However, the accuracy of weather-yield relationships derived from crop models, whether statistical or process-based, is dependent on the quality of the underlying input data used to run these models. In this context, a major challenge in many developing countries is the lack of accessible and reliable meteorological datasets. Gridded weather datasets, derived from combinations of *in situ* gauges, remote sensing, and climate models, provide a solution to fill this gap, and have been widely used to evaluate climate impacts on agriculture in data-scarce regions worldwide. However, these reference datasets are also known to contain important biases and uncertainties. To date, there has been little research to assess how the choice of reference datasets influences projected sensitivity of crop yields to weather. We compare multiple freely available gridded datasets that provide daily weather data over the Indian sub-continent over the period 1983–2005, and explore their implications for estimates of yield responses to weather variability for key crops grown in the region (wheat and rice). Our results show that individual gridded weather datasets vary in their representation of historic spatial and temporal temperature and precipitation patterns across India. We show that these differences create large uncertainties in estimated crop yield responses and exposure to variability in growing season weather, which in turn, highlights the need for improved consideration of input data uncertainty in statistical studies that explore impacts of climate variability and change on agriculture.

1. Introduction

Farmer livelihoods depend strongly on weather conditions during the growing season. Smallholder subsistence farmers in developing countries in Africa and Asia, in particular, are impacted disproportionately by weather shocks due to their lower adaptive capacity and dependence on agriculture for basic staple crop production, nutrition and incomes (Morton 2007, Harvey *et al* 2014, Niles and Salerno 2018). In this

context, understanding crop yield responses in smallholder farming systems to different types and magnitudes of weather shocks is critical for estimating impacts of future climate variability and change on agricultural productivity and food security, and for designing appropriate strategies to reduce exposure to weather-related production risks.

To assess the impacts of weather and future climate on agriculture, crop models are commonly used to simulate yield responses to different meteorological

conditions. Two types of modelling approaches exist. Statistical yield models (e.g. Lobell and Burke 2010, Cai *et al* 2014, Duncan *et al* 2015, 2016, Parkes *et al* 2017, 2018) develop empirical relationships between observed weather conditions and crop yields reported through field surveys or agricultural censuses. In contrast, process-based models, such as APSIM (Holzworth *et al* 2014) or AquaCrop (Foster *et al* 2017), use mathematical representations of plant physiology to simulate crop growth and yield development for specified meteorological conditions, soil properties, and management practices. Both modelling approaches have strengths and weaknesses (Roberts *et al* 2017). Importantly, where sufficient observed yield data exist, statistical models may provide additional information about yield sensitivity to climate due to their ability to account for the effects of unobserved farmer management practices or indirect weather-related drivers of yield losses that cannot be simulated by process-based models (e.g. mechanical damage by hail or wind, pests, diseases, extreme rainfall etc) (Roberts *et al* 2017, Li *et al* 2019).

In many regions worldwide, and in particular in developing countries where agriculture underpins food security and rural livelihoods, there is a lack of reliable and comprehensive historical weather records from *in situ* monitoring stations. Consequently, statistical models typically are developed using weather data drawn from national, regional, or global gridded weather datasets. A diverse range of gridded weather datasets exist (e.g. Yatagai *et al* 2012, Ashouri *et al* 2015, Funk *et al* 2015, Ruane *et al* 2015), each differing in the underlying source of primary observations (e.g. satellite data, model reanalysis, etc), the variables reported (e.g. temperature, precipitation, solar radiation, etc), and the resolution (spatial and temporal) at which these data are reported. Differences in the data sources and algorithms used to create these gridded products mean that reported meteorological conditions at a given location and time can often vary substantially across datasets. For example, the high resolution of 0.05° used in the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) (Funk *et al* 2015) is able to resolve storms a few kilometres across which are blurred out by the coarser resolution datasets such as the 0.75° ECMWF Re-Analysis Interim (ERA-Interim) (Dee *et al* 2011).

Despite these known differences, the choice of weather dataset is often an arbitrary decision in studies that use these datasets to evaluate climate impacts on agriculture and other sectors (Auffhammer *et al* 2012, Cai *et al* 2014, Duncan *et al* 2016). To date, there has been little evaluation of how the choice of reference weather datasets affects estimates of implied sensitivity of agriculture to climate variability and change from statistical crop yield models. This omission is in contrast with hydrological modelling (Decharme and Douville 2006) and statistical climate downscaling (Iizumi *et al* 2017) where weather dataset choice has been acknowledged as a key source of model uncertainty. Similarly, higher

spatial resolution datasets may be able to better to capture elevation-induced weather variability, which is known to be a key driver of uncertainty in gridded weather products (Semenov *et al* 2013, Beck *et al* 2019).

In this study, we address this knowledge gap by developing multiple statistical crop yield models for wheat and rice production in India using 50 unique combinations of weather datasets. Our study focuses on India due to the availability of comprehensive historical yield observation data, along with the widespread past use of statistical modelling approaches in India and South Asia as part of climate impact research (Lobell *et al* 2012, Duncan *et al* 2015, 2016, Mondal *et al* 2016, Asseng *et al* 2017, Jain *et al* 2017, Gilmont *et al* 2018). We hypothesise that weather dataset choice will have an impact on estimates of crop yield responses to weather shocks. These shocks, in turn, will introduce uncertainty to estimates of farmers responses to weather-related production risks and impacts of future climate change on agriculture.

2. Methods

2.1. Study area
Crop yield observations required to train statistical models were obtained from a panel dataset of district-level yield observations for rice and wheat from across India provided by the ICRISAT VDSA (Village Dynamics in South Asia) study database (http://vdsa. icrisat.ac.in/vdsa-database.aspx). For this study, we use yield data for the period 1983–2005 to align with the common temporal coverage period for available gridded weather products (table 1), and omit district-year observations where reported cropped area of wheat or rice was less than 1000 acres in any year of our record. These restrictions retain 267 districts for wheat and 299 districts for rice. Districts for each crop are geographically spread across most of India, omitting mostly the extreme northern districts where terrain is mountainous, along with a small number of districts in eastern India, where administrative boundary changes make identifying consistent locations infeasible.

2.2. Weather datasets
To develop statistical models of weather impacts on wheat and rice yields, we use a total of 50 different reference gridded weather dataset combinations (precipitation and temperature) that are available for the Indian sub-continent. Table 1 summarises the key features of these datasets (name, resolution, variables, data sources, and key references). Almost all districts in India are significantly larger than the spatial resolution of gridded weather products. As a result, seasonal temperature and precipitation variables are calculated for each pixel and subsequently aggregated up to the district-level to match the spatial resolution of crop yield observations. Aggregation is performed based on area weighted averages of weather dataset

Table 1. Name, abbreviation, resolution, indication of temperature or precipitation data, time frame, primary type of source data and key reference for the reference datasets. Where *T* and *P* indicate if the dataset provides temperature or precipitation data respectively. Datasets that list 'Present' as their end time can be delayed by a few months. See S6 for a visual representation of the differences in resolution.

Name	Abbrev.	Res.	*T*	*P*	Time frame	Primary source type	References
AgMERRA	Ag	0.25°	X	X	1980–2010	Bias corrected reanalysis	Ruane *et al* (2015)
APHRODITE	AP	0.50°		X	1951–2007	Weather station data	Yatagai *et al* (2012)
CHIRPS	CH	0.05°		X	1981—Present	Bias corrected satellite record	Funk *et al* (2015)
ERA-Interim	ER	0.75°	X	X	1979—Present	Reanalysis	Dee *et al* (2011)
IMD-Precipitation	IM	0.50°		X	1969–2005	Weather station data	Pai *et al* (2014), Rajeevan and Bhate (2009)
IMD-Temperature	IM	1.00°	X		1969–2005	Weather station data	Pai *et al* (2014), Rajeevan and Bhate (2009)
PERSIANN-CDR	PE	0.25°		X	1983—Present	Satellite record	Ashouri *et al* (2015)
POWER	PO	0.50°	X	X	1981—Present	Satellite record	Stackhouse *et al* (2018)
Princeton	Pr	0.25°	X	X	1948–2016	Bias corrected reanalysis	Sheffield *et al* (2006)
WFDEI-CRU	WC	0.50°		X	1979–2016	Bias corrected reanalysis	Weedon *et al* (2014), Harris *et al* (2013)
WFDEI-GPCC	WG	0.50°		X	1979–2016	Bias corrected reanalysis	Weedon *et al* (2014), Schneider *et al* (2014)

pixels within each fixed district boundary, as opposed to specific annual harvested areas for which no reliable inter-annual data exists for our spatial and temporal study domain. Furthermore, we do not consider differences in elevation between grid cells when weighting, consistent with prior statistical crop modelling (Fishman 2016, Zaveri and Lobell 2019) when upscaling gridded weather data.

Of the gridded weather datasets included in our analyses (table 1), three (some reporting both temperature and precipitation) rely on satellite data as a primary source of weather observations—PERSIANN-CDR, NASA POWER, and CHIRPS. In contrast, two datasets use weather station data as their primary source of information for generating gridded weather observations—the APHRODITE, and Indian Meteorological Department (IMD) datasets. The final five datasets use weather model reanalysis as their data source, whether bias corrected or uncorrected: AgMERRA, ERA-Interim, Princeton, and two variants of the WATCH-Forcing-Data-ERA-Interim (WFDEI-CRU and WFDEI-GPCC) that are differentiated based on the reference precipitation dataset used to generate the product. The sources here are not necessary the only source used and datasets such as CHIRPS use a blend of satellite and station data that is subsequently bias corrected. See S1 and S2 for examples of the differences between the datasets.

Many of the weather datasets reported in table 1 have been used in past studies of climate impacts on agriculture in India and more broadly in South Asia. For example, the precipitation only datasets (i.e. CHIRPS, APHRODITE, PERSIANN-CDR, IMD) have been used to analyse impacts of drought and rainfall extremes on agricultural yields and water demands in the region (Romaguera *et al* 2010, DeFries *et al* 2016, Aadhar and Mishra 2017). Similarly, the Princeton and ERA-Interim datasets have been used in

assessing regional droughts (Mishra *et al* 2014) and water resource assessments (Mathison *et al* 2013), while WFDEI has been used as an input in a study on irrigation demand (Biemans *et al* 2016). AgMERRA has been used to analyse the uncertainty of aggregating crop yields from large scale studies (Porwollik *et al* 2017) and for a global gridded crop model evaluation (Müller *et al* 2017). Finally, NASA POWER has been widely applied for modelling temperature—related impacts on wheat yields in the South Asia region (Asseng *et al* 2017). Although not an exhaustive list, these studies highlight both the regional relevance of the datasets and their history of use in agricultural and climate impacts research across the Indian sub-continent. Additional global gridded weather datasets, for example AgCFSR (Ruane *et al* 2015) and S14FD (Iizumi *et al* 2017) that have been specifically developed for crop modelling, also exist and should be considered in future research beyond the Indian sub-continent

2.3. Yield models

We develop a total of 50 unique statistical crop yield models for both rice and wheat in India, using the different combinations of available gridded temperature and precipitation datasets described in table 1 and section 2.2. Following Lobell and Burke (2010), wheat yield models are formulated as linear regressions relating observed crop yields to growing season aggregates of precipitation and temperature (growing degree days and extreme degree days). The rice yield model uses a similar specification to the wheat yield models, but omits the extreme degree day term to maintain consistency with prior statistical models of rice yields in South Asia (Auffhammer *et al* 2012, Fishman 2016). Specifically, the wheat and rice yield models are defined as follows:

Wheat: $\ln(Y_{t,i}) = bGDD_{t,i} + cEDD_{t,i} + dP_{t,i}$
$$+ f_s(t) + a_i + \epsilon_{t,i} \quad (1)$$

Rice: $\ln(Y_{t,i}) = bGDD_{t,i} + dP_{t,i} + f_s(t) + a_i + \epsilon_{t,i}$
$$(2)$$

where: $Y_{t,i}$ is the crop yield from district i in year t, $GDD_{t,i}$ is the total seasonal growing degree days, $EDD_{t,i}$ is the total seasonal extreme degree days (EDD), $P_{t,i}$ is the total seasonal precipitation, b, c and d are model parameters, $f_s(t)$ is a state-specific quadratic time-trend representing growth in yields through breeding and other improvements to management practices, a_i is a district-specific fixed effect term (i.e. district-level intercept) to account for unobserved time-invariant drivers of yield differences between districts, and $\epsilon_{t,i}$ is the error term. The natural log of yield is taken to produce models that provide relative changes instead of absolute changes, since this moderates the effects of districts having significantly different absolute yields.

Precipitation, growing degree days (GDD) and extreme degree days per season are calculated as the sum of the daily values of these variables within the respective growing seasons for wheat (November to February) and rice (June to September) (Datta and Jong 2002, Auffhammer *et al* 2012). Daily values of GDD and EDD are calculated as shown in equations (3) and (4) below, accounting for the within-day distribution of temperatures by fitting a sinusoidal curve between observed maximum and minimum temperatures on each day following the approach proposed by Schlenker and Roberts (2009). This approach to GDD and EDD estimation is selected as it is provides a more robust estimate of daily degree day accumulation compared to using a simple daily average temperature, which does not account for within-day temperature distributions and thus may affect statistical model performance (Fontes *et al* 2017, Roberts *et al* 2017).

$$GDD = \int_{T_{base}}^{\infty} \min(T - T_{base}, T_{upp} - T_{base}) \, \phi(T)$$
$$(3)$$

$$EDD = \int_{T_{str}}^{\infty} \max(T - T_{str}, 0) \, \phi(T), \quad (4)$$

where T is the temperature, $\phi(T)$ is the daily cumulative distribution of interpolated temperatures based on the sinusoidal fit between the daily maximum and minimum temperatures, and T_{base} and T_{upp} are crop-specific lower and upper temperature limits, respectively, for GDD accumulation. For wheat, these lower and upper limits are equal to 0 °C and 30 °C (Lobell *et al* 2012). In contrast, for rice, limits of 8 °C and 30 °C are chosen following van Oort *et al* (2011). EDD accumulation occurs for temperatures above a threshold temperature limit for the initiation of heat stress (T_{str}), which we set equal to 30 °C for wheat consistent with prior econometric yield models (Schlenker and Roberts 2009, Tack *et al* 2015, Roberts *et al* 2017). As

described earlier, no value is specified for rice as EDDs are omitted from these models.

2.4. Assessing uncertainty in future crop yield change

A common use of statistical crop yields models is to assess how agricultural production may respond to future changes in climate conditional on existing management practices and technologies being hold constant (Lobell and Burke 2010). In this study, we explore the effect of the choice of the reference weather training dataset on the estimated impacts of climate change on Indian agriculture by applying a range of idealized future temperature and precipitation change scenarios to our set of 50 unique statistical wheat and rice yield models. Temperature change scenarios in our analysis involved perturbing baseline daily temperature values by between −2 °C and +2 °C in increments of 1 °C in each reference dataset. Updated seasonal GDD and EDD totals in each district and year are then calculated using these perturbed daily temperature time series holding threshold limits for GDD/EDD accumulation constant consistent with assumptions of no changes in crop varietal properties in this analysis. Precipitation change scenarios in turn were constructed by modifying existing seasonal precipitation totals for each dataset, district and year by between −20% and +20% in steps of 10%. We apply each combination of perturbed precipitation and temperature (GDD and EDD) totals as inputs to statistical yield models generated in section 2.3 to generate projections of future yield changes, and, in particular, evaluate uncertainty in yield change projections resulting from the choice of historical reference weather dataset.

3. Results

Our results show that only limited differences exist in the ability of statistical models developed using different reference weather datasets to explain observed spatial and temporal variability in observed crop yields. All models are highly significant ($p < 0.001$), and capture a large proportion of the observed spatial and temporal variability in district-level wheat ($R^2 > 0.85$) and rice ($R^2 > 0.80$) yields. Performance of models is shown to be robust when assessed based out-of-sample tests omitting one-year of data at a time (SI figure 7 is available online at stacks. iop.org/ERL/14/124089/mmedia), with only limited differences observed between in-sample and out-of-sample correlations with actual yields across all dataset combinations. However, the choice of reference weather dataset does lead to large differences in the significance and magnitude of individual estimated weather coefficients. From this, several key insights can be drawn to inform the use of such models in
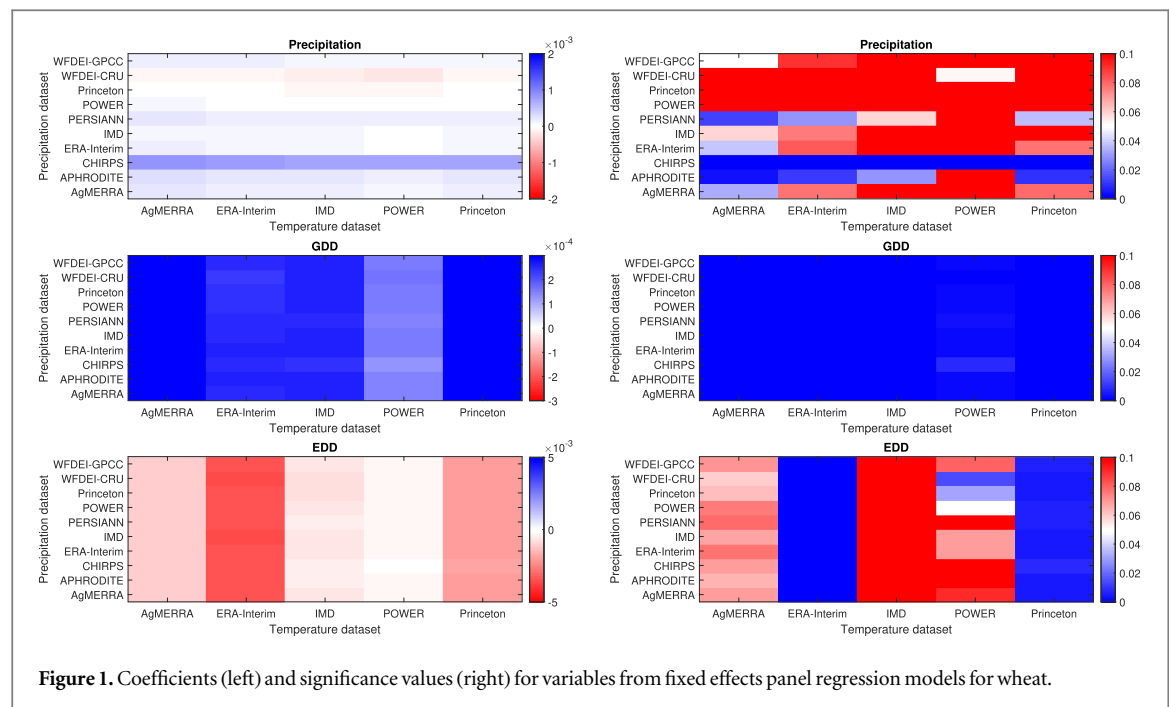
Figure 1. Coefficients (left) and significance values (right) for variables from fixed effects panel regression models for wheat.
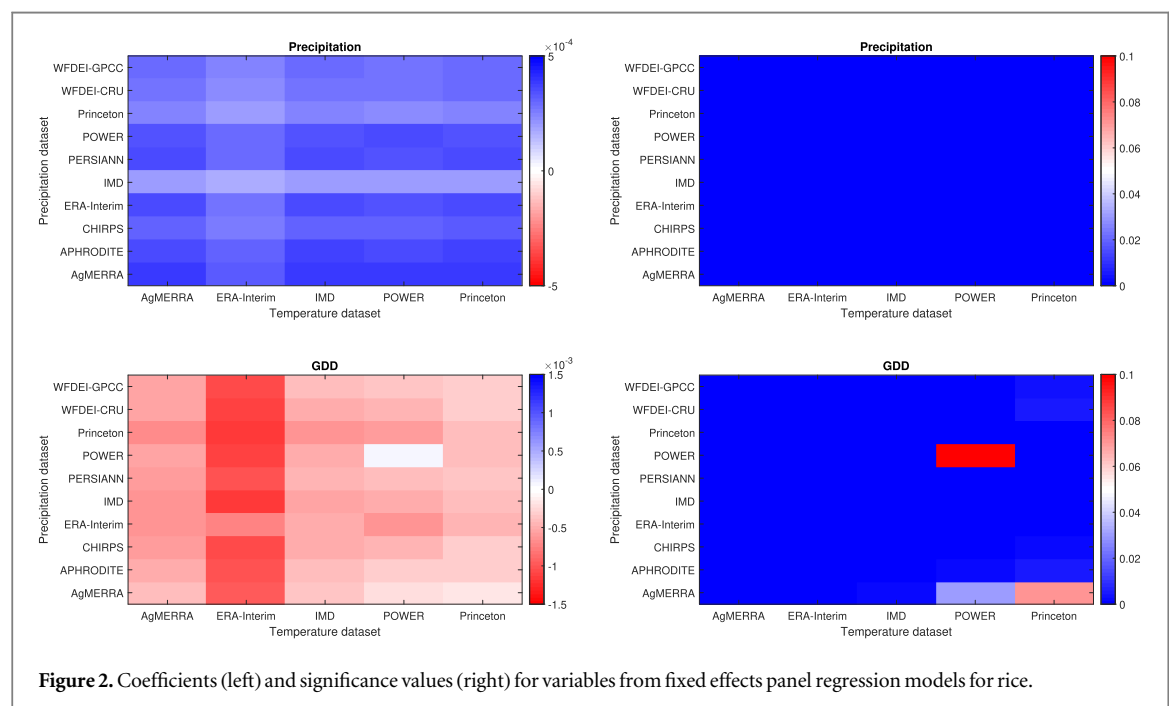
Figure 2. Coefficients (left) and significance values (right) for variables from fixed effects panel regression models for rice.

weather and climate impact assessment (figures 1 and 2).

3.1. Weather dataset choice alters implied crop yield sensitivity to climate

While all models perform equally well in explaining overall yield variability, figures 1 and 2 highlight that only limited consensus exists across models about the magnitude and significance of crop yield sensitivity to specific meteorological variables and extreme events. For wheat, GDDs are a significant ($p < 0.05$) and positive predictor of crop yields (i.e. yields increase with GDDs) for all datasets, in agreement with

previous studies of wheat production in South Asia (Mondal *et al* 2016). Increasing EDDs has a negative impact on wheat yields for all datasets. However, the negative impact of EDD's is only statistically significant ($p < 0.05$) for two datasets based on the temperature thresholds and specifications adopted in this analysis, highlighting that weather dataset choice can have important implications for robustness of conclusions drawn about climatic driver of yield variability. A clear difference in temperature coefficients—both for GDD and EDD—is also observed for models using NASA POWER temperature data, for which coefficient sizes are noticeably smaller than other

temperature datasets (figure 1). This can be explained by the hot bias in POWER temperature data relative to other datasets (figure S1), resulting in higher EDD totals and therefore a smaller coefficient. POWER's hot temperature bias also has the effect of blurring the identification of positive GDD and negative EDD effects, as the temperature threshold for wheat is assumed to be crop rather than dataset specific. Figure S1 demonstrates this effect, showing that more than 65% of daily observations during the wheat growing season have a higher maximum temperature than 30 °C—the lower limit for EDD accumulation—for POWER. By contrast, ERA-Interim has the smallest percentage of days (14%) that exceed 30 °C during the wheat growing season. We suggest that this may explain the greater significance and larger coefficient size for EDD in wheat models using ERA-Interim temperature data, which captures a much smaller subset of true extreme temperature events in comparison with other datasets.

In general, higher precipitation has a positive effect on wheat yields for all datasets except WFDEI-CRU. However, for the majority of datasets, precipitation is not a significant predictor ($p > 0.05$) of wheat yields, reflecting the fact that wheat is commonly irrigated across much of India. Indeed, it is noticeable that for many datasets the identification of precipitation as a predictor of yield variability is not robust, with significance changing substantially depending on the choice of paired temperature dataset. CHIRPS is the only dataset for which precipitation has a consistently significant relationship with wheat yields. CHIRPS predicts a stronger effect than any of the other precipitation datasets, perhaps reflecting the greater capacity of CHIRPS to capture aggregated impacts of sub-district rainfall heterogeneity. However, it is important to note that the absolute differences in the size of precipitation coefficients are small across datasets (figure S3). A 10 mm change in total seasonal precipitation is between 15% and 30% of the total seasonal precipitation (figure S2). Yield differences for a precipitation change of this magnitude range between 14.2 kg/ha for CHIRPS to −2.2 kg/ha for WFDEI-GPCC, both of which are less than 1% of the average wheat yield in the dataset.

For rice, precipitation is a consistently positive and significant predictor of variability in crop yields over space and time in India (figure 2), reflecting the fact that rice is predominantly grown under rainfed or partial irrigation conditions (DeFries *et al* 2016). However, variations in the size of the precipitation coefficients exist between datasets, with coefficient values ranging from 1.66×10^{-4} to 3.90×10^{-4} ln(kg/ha)/mm. These values mean that an increase in seasonal rainfall of 100mm would result in a yield increase of between 31.4 kg/ha and 62.2 kg/ha, a non-trivial level of uncertainty as total monsoon (kharif) seasonal precipitation (averaged across all years and datasets) in India is 869 mm. For all datasets,
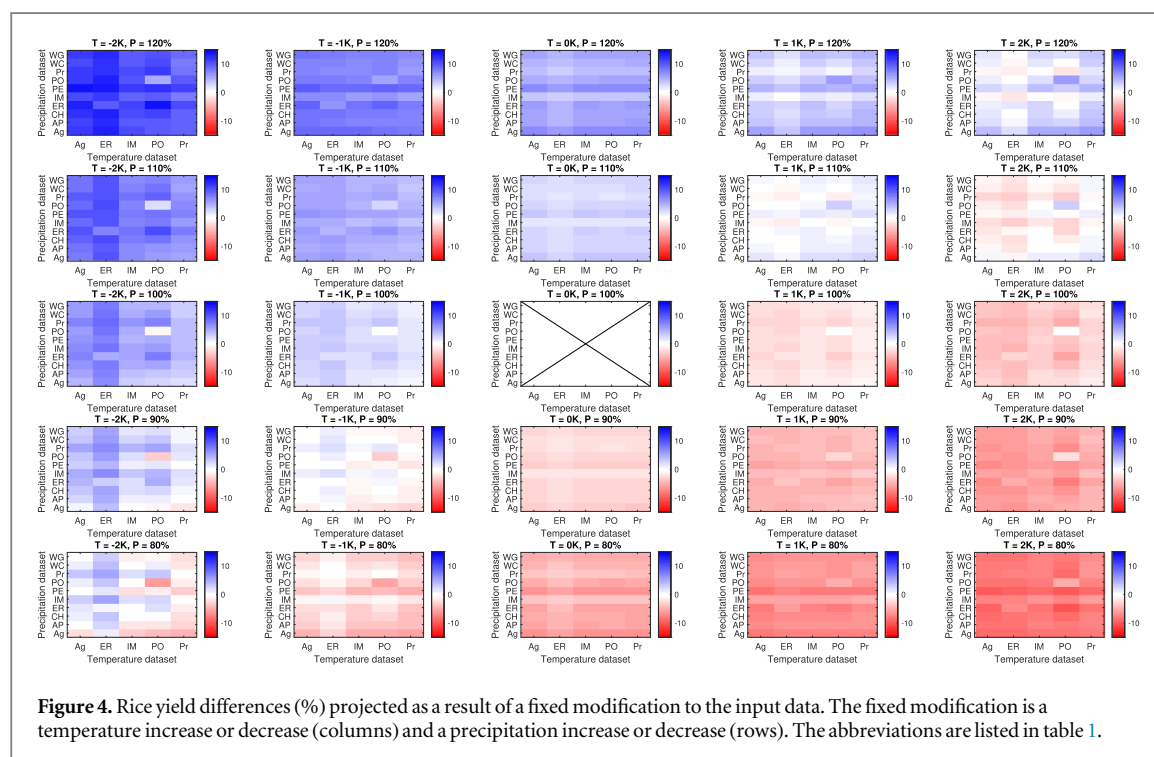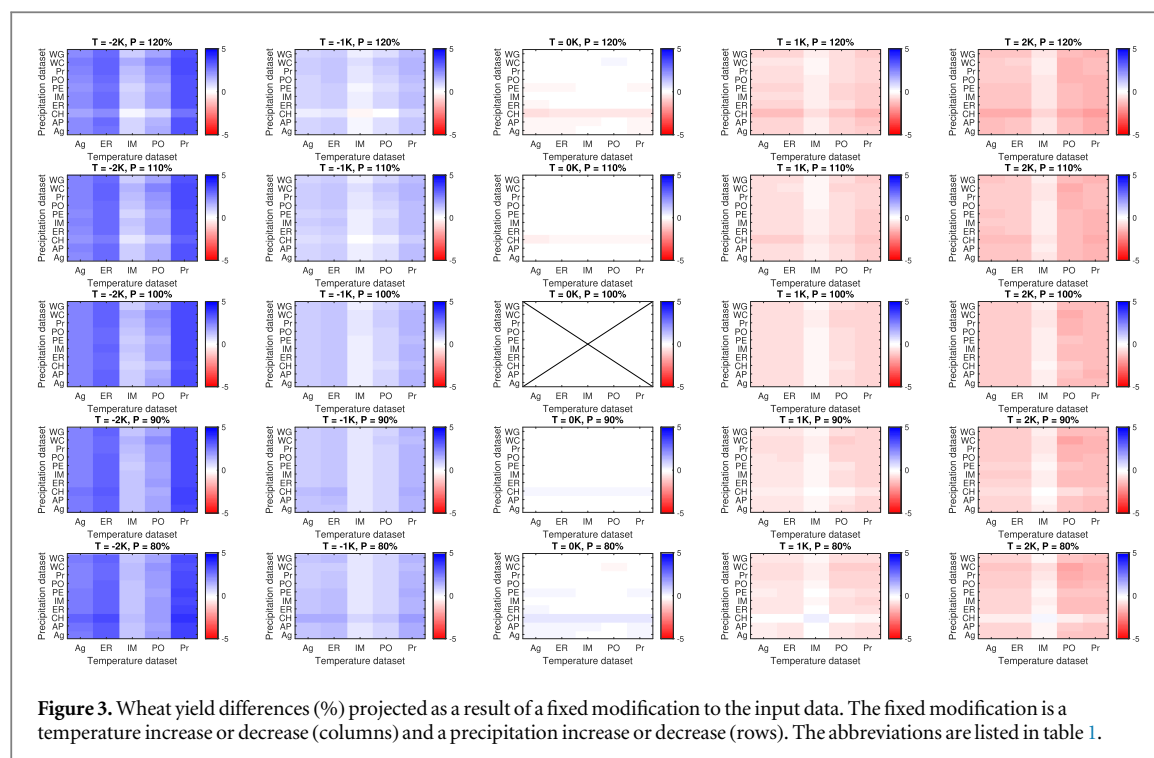
GDDs are found to have a negative impact on rice yields. This finding indicates that higher temperatures lead to yield reductions, and is consistent with previous econometric studies of rice yields in India (Auffhammer *et al* 2012, Fishman 2016). However, as with precipitation, notable heterogeneity exists across datasets, with yield reductions per 100 GDDs ranging from 56.3 kg/ha for Princeton to 177.1 kg/ha when using ERA-Interim.

The models in this study are based on seasonal totals for precipitation, GDDs and EDDs instead of the sub-seasonal totals used with most process-based models. Crop yields are known to be affected by intra-seasonal timing (Dalhaus *et al* 2018, Hufkens *et al* 2019) and intensity (Fishman 2016) of extreme weather events. Incorporating sub-seasonal weather variables is likely to exacerbate differences between models even further as individual gridded weather datasets are known to differ substantially in their ability to capture intraseasonal weather dynamics, for example the size and arrival of the South Asian monsoon (Ceglar *et al* 2017).

3.2. Predicted impacts of climate change vary with reference dataset choice

The changes in predicted wheat yields for each model under potential climate change scenarios are shown in figure 3. The range in yield changes for a one degree increase in temperature is between −0.45% ± 0.07% when using models trained on IMD temperature data, and −1.15% ± 0.01% when using models trained on ERA-Interim temperature data (where the uncertainty is the standard deviation across the precipitation datasets for a given temperature dataset). For a two degree increase in temperature, yield changes for wheat expand further to −1.13% ± 0.18% (IMD dataset) and −2.97% ± 0.03% (ERA-Interim dataset). These two results highlight how the low temperature bias in ERA-Interim affects the yield-temperature relationship for projected crop yields. The increase in temperature leads to an increase in EDDs, which in turn reduce the yield. Effects of precipitation changes on yields are smaller in magnitude, with the changes in yields for a 20% increase in seasonal precipitation ranging from −0.10% ± 0.04% (WFDEI-GPCC dataset) to 0.59% ± 0.28% (CHIRPS dataset) reflecting the smaller effects of precipitation on wheat yields in India.

Results of the climate scenario analyses for rice are shown in figure 4. Rice results show a significantly larger disagreement in future yield changes between datasets, reflecting greater heterogeneity in coefficient sizes for baseline rice models as shown in figure 2. For example, a 20% increase in seasonal precipitation for rice leads to an increase in yields of between 3.47% ± 0.26% (IMD dataset) and 6.67% ± 0.53% (AgMERRA dataset). In contrast, a two degree increase in temperature results in rice yield reductions of
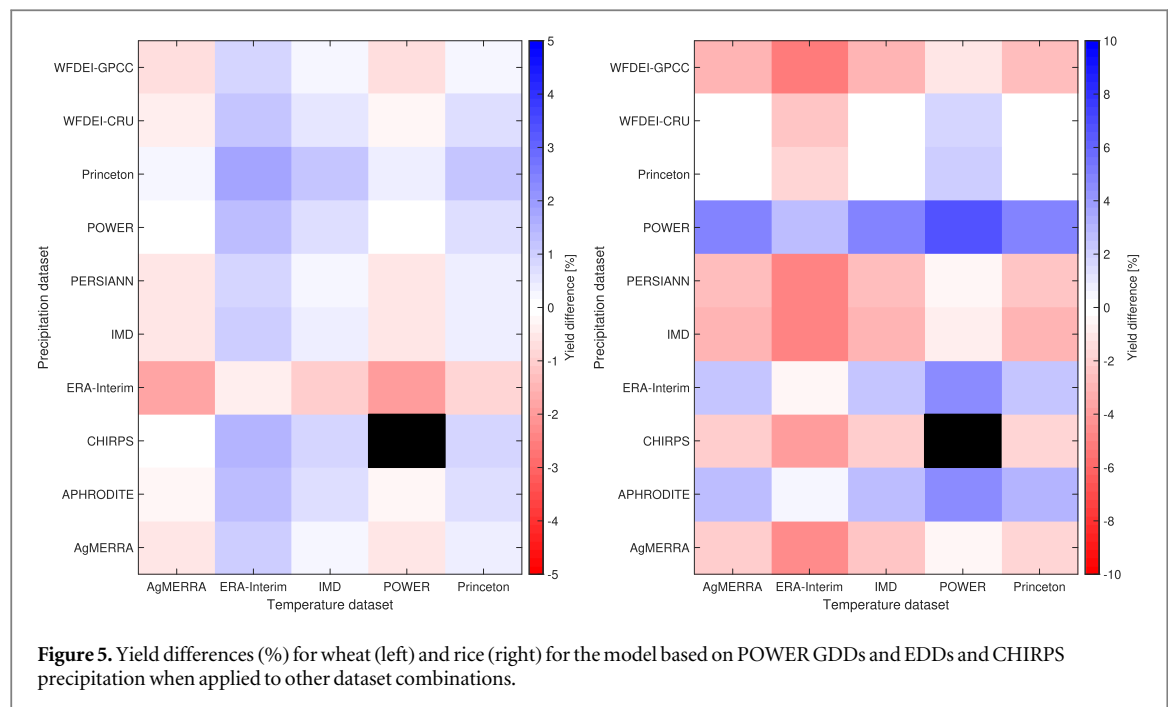
Figure 3. Wheat yield differences (%) projected as a result of a fixed modification to the input data. The fixed modification is a temperature increase or decrease (columns) and a precipitation increase or decrease (rows). The abbreviations are listed in table 1.

Figure 4. Rice yield differences (%) projected as a result of a fixed modification to the input data. The fixed modification is a temperature increase or decrease (columns) and a precipitation increase or decrease (rows). The abbreviations are listed in table 1.

anywhere between 3.50% ± 0.76% (Princeton dataset) and 7.51% ± 0.98% (ERA-Interim dataset) when holding precipitation constant at historical levels. Notably, there is large uncertainty in the combined effects of uncertain future changes in temperature and precipitation. For example, yield changes for a one degree temperature increase and a 10% precipitation increase range from 2.46% for the POWER temperature + POWER precipitation model to 6.32% for POWER temperature + ERA-Interim precipitation

model—a spread of over 3% based on reference dataset choice alone.

4. Discussion and implications

Understanding the meteorological drivers of crop yield variability is important for assessing exposure of agriculture to climate risks, and for designing effective strategies to mitigate impacts of future climate change. For a case study in India, our findings highlight that

Figure 5. Yield differences (%) for wheat (left) and rice (right) for the model based on POWER GDDs and EDDs and CHIRPS precipitation when applied to other dataset combinations.

our understanding of the sensitivity of crop yields to climate variability and change is affected strongly by the choice of reference weather dataset used to train baseline statistical models. These differences in turn have important implications for estimation of impacts of climate change on agriculture (Van Wart *et al* 2013), which should be considered in future research alongside other sources of uncertainty in climate impact assessments such as the choice of emission scenarios, climate model specification, downscaling approach, reference timescale or impact model (Oettli *et al* 2011, Hawkins *et al* 2013, Chen *et al* 2015, Iizumi *et al* 2017).

In order to address these challenges, there is an urgent need for greater evidence about the ability of different gridded data products to capture spatial and temporal weather variability in major agricultural regions. A number of studies have evaluated the performance of various gridded weather datasets against station-level observations globally and regionally (Iizumi *et al* 2014, Ruane *et al* 2015, Behnke *et al* 2016, Beck *et al* 2017, Mourtzinis *et al* 2017). However, conclusions drawn from these studies are primarily driven by performance in areas with high densities of weather stations with long-term records (e.g. North America, Europe). In contrast, performance of gridded weather products in smallholder farming environments in Africa and Asia is less well understood and quantified, due to the more limited availability, coverage, and reliability of weather station data in these locations (Menne *et al* 2012, Van Wart *et al* 2015, Heft-Neal *et al* 2017). In these regions, our findings suggest that modellers therefore should be cautious in using only a single gridded weather data product to understand current and future agricultural climate risks. Specifically, we recommend that multiple gridded

weather datasets should instead be used when developing statistical crop yield models in absence of information about the most reliable gridded weather dataset, an approach that is comparable to the use of multi-model ensembles in climate and other geophysical modelling studies (Tebaldi and Knutti 2007, Rosenzweig *et al* 2014).

Alongside these recommendations, our findings also highlight the importance of weather dataset consistency throughout the design and application of statistical weather-yield models. As an illustrative example, figure 5 shows the errors in estimated average wheat and rice yields when a statistical model is trained on POWER temperature and CHIRPS precipitation dataset and then used to predict yields using alternative combinations of different input precipitation and temperature datasets (see figures S4 and S5 for results using all dataset combinations). Importantly, such errors may have significant implications for several end uses of statistical crop yield models. For example, weather index insurance policies, which are widely offered to smallholder farmers in India and other regions as a way to help mitigate financial risks posed by weather-related crop losses (Barnett and Mahul 2007, Clarke *et al* 2012), are often designed and implemented using a range of different weather data sources (e.g. long-term gridded weather data for regional contract design versus short-term station data for triggering localised payouts). Each of these may contain a different underlying bias. Where differences in biases are large, the performance of index insurance products may be negatively affected due to an over-estimation or under-estimation of underlying climate risks for farmers. This insight is comparable with previous conclusions about the robustness of weather index insurance under non-stationary climate, which

highlight that insurance performance deteriorates as weather conditions deviate from historical benchmarks due to factors such as multi-decadal climate variability and man-made climate change (Daron and Stainforth 2014).

Finally, while the discussion thus far in this paper has focused on the impacts of gridded weather dataset differences in the context of statistical crop yield modelling, it is important to note that dataset bias will also pose similar challenges when using biophysical process-based crop models. Process-based crop models internally specify fixed biological relationships between growing season weather conditions and crop yields. For example, the APSIM model assumes that terminal heat stress for wheat is initiated for temperatures above 34 °C. In addition, cold biases in weather input data could also lead to an under-estimation of yields because due to the erroneous triggering of low temperature stress occurred within a process-based crop model (Iizumi *et al* 2010). Any biases in input weather datasets therefore will alter process-based model predictions of expected yield variability, which, in turn, may result in either an over- or under-estimation of weather-related production risks in a given farming system. The use of $T_{str} = 30$ °C is a limiting factor in this study as it is based on values from other statistical analyses (Schlenker and Roberts 2009, Tack *et al* 2015, Roberts *et al* 2017) and the datasets therein. The use of fixed limits is an additional mechanism through which weather dataset choice can create uncertainty in statistical yield models. For example using an input weather dataset with a hot temperature bias (e.g. POWER) will, all else being equal, lead a process-based crop model to predict greater frequencies and magnitudes of yield losses due to extreme heat than if the same simulations were run using an input temperature dataset without such a bias. As with statistical crop yield modelling, in the absence of objective information about the accuracy of different gridded weather datasets, we argue that addressing this challenge requires greater use of ensembles of gridded weather datasets in process-based crop model simulations. Such an approach would provide a more accurate picture of the uncertainty in estimates of the exposure of agriculture to climate risks, and, in turn, improve the robustness of policy and management recommendations about how to improve resilience of smallholder farming to extreme weather and climate change.

5. Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

ORCID iDs

B Parkes ⓘ https://orcid.org/0000-0002-5936-8331
T P Higginbottom ⓘ https://orcid.org/0000-0003-1150-1444
K Hufkens ⓘ https://orcid.org/0000-0002-5070-8109
F Ceballos ⓘ https://orcid.org/0000-0001-8699-5114
B Kramer ⓘ https://orcid.org/0000-0001-7644-6613
T Foster ⓘ https://orcid.org/0000-0001-9594-1556

References

Aadhar S and Mishra V 2017 High-resolution near real-time drought monitoring in south Asia *Sci. Data* **4** 170145

Ashouri H, Hsu K-L, Sorooshian S, Braithwaite D K, Knapp K R, Cecil L D, Nelson B R and Prat O P 2015 Persiann-cdr: daily precipitation climate data record from multisatellite observations for hydrological and climate studies *Bull. Am. Meteorol. Soc.* **96** 69–83

Asseng S, Cammarano D, Basso B, Chung U, Alderman P D, Sonder K, Reynolds M and Lobell D B 2017 Hot spots of wheat yield decline with rising temperatures *Glob. Change Biol.* **23** 2464–72

Auffhammer M, Ramanathan V and Vincent J R 2012 Climate change, the monsoon, and rice yield in India *Clim. Change* **111** 411–24

Barnett B J and Mahul O 2007 Weather index insurance for agriculture and rural areas in lower-income countries *Am. J. Agric. Econ.* **89** 1241–7

Beck H E, Vergopolan N, Pan M, Levizzani V, van Dijk A I, Weedon G P, Brocca L, Pappenberger F, Huffman G J and Wood E F 2017 Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling *Hydrol. Earth Syst. Sci.* **21** 6201–17

Beck H E, Wood E F, McVicar T R, Zambrano-Bigiarini M, Alvarez-Garreton C, Baez-Villanueva O M, Sheffield J and Karger D N 2019 Bias correction of global high-resolution precipitation climatologies using streamflow observations from 9372 catchments *J. Clim.* (https://doi.org/10.1175/JCLI-D-19-0332.1)

Behnke R, Vavrus S, Allstadt A, Albright T, Thogmartin W E and Radeloff V C 2016 Evaluation of downscaled, gridded climate data for the conterminous United States *Ecol. Appl.* **26** 1338–51

Biemans H, Siderius C, Mishra A and Ahmad B 2016 Crop-specific seasonal estimates of irrigation-water demand in south Asia *Hydrol. Earth Syst. Sci.* **20** 1971–82

Cai R, Yu D and Oppenheimer M 2014 Estimating the spatially varying responses of corn yields to weather variations using geographically weighted panel regression *J. Agric. Resour. Econ.* **39** 230–52

Ceglar A, Toreti A, Balsamo G and Kobayashi S 2017 Precipitation over monsoon asia: a comparison of reanalyses and observations *J. Clim.* **30** 465–76

Chen J, Brissette F P and Lucas-Picher P 2015 Assessing the limits of bias-correcting climate model outputs for climate change impact studies *J. Geophys. Res.: Atmos.* **120** 1123–36

Clarke D J, Mahul O, Rao K N and Verma N 2012 Weather based crop insurance in India *Policy Research Working Paper 5985* (Washington, DC: The World Bank) (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2016750)

Dalhaus T, Musshoff O and Finger R 2018 Phenology information contributes to reduce temporal basis risk in agricultural weather index insurance *Sci. Rep.* **8** 46

Daron J D and Stainforth D A 2014 Assessing pricing assumptions for weather index insurance in a changing climate *Clim. Risk Manage.* **1** 76–91

Datta K and Jong C 2002 Adverse effect of waterlogging and soil salinity on crop and land productivity in northwest region of Haryana, India *Agric. Water Manage.* **57** 223–38

Decharme B and Douville H 2006 Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations *Clim. Dyn.* **27** 695–713

Dee D P *et al* 2011 The era-interim reanalysis: configuration and performance of the data assimilation system *Q. J. R. Meteorol. Soc.* **137** 553–97

DeFries R, Mondal P, Singh D, Agrawal I, Fanzo J, Remans R and Wood S 2016 Synergies and trade-offs for sustainable agriculture: nutritional yields and climate-resilience for cereal crops in central India *2nd Int. Glob. Food Security Conf.: Glob. Food Secur.* **11** 44–53

Duncan J, Saikia S, Gupta N and Biggs E 2016 Observing climate impacts on tea yield in Assam, India *Appl. Geogr.* **77** 64–71

Duncan J M A, Dash J and Atkinson P M 2015 Elucidating the impact of temperature variability and extremes on cereal croplands through remote sensing *Glob. Change Biol.* **21** 1541–51

Fishman R 2016 More uneven distributions overturn benefits of higher precipitation for crop yields *Environ. Res. Lett.* **11** 024004

Fontes F P, Gorst A and Palmer C 2017 Does choice of drought index influence estimates of drought-induced cereal losses in India? *GRI Working Papers 274* (Grantham Research Institute on Climate Change and the Environment)

Foster T, Brozović N, Butler A, Neale C, Raes D, Steduto P, Fereres E and Hsiao T C 2017 AquaCrop-OS: an open source version of FAO's crop water productivity model *Agric. Water Manage.* **181** 18–22

Funk C *et al* 2015 The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes *Sci. Data* **2** 150066

Gilmont M, Hall J, Grey D, Dadson S, Abele S and Simpson M 2018 Analysis of the relationship between rainfall and economic growth in Indian states *Glob. Environ. Change* **49** 56–72

Harris I, Jones P, Osborn T and Lister D 2013 Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset *Int. J. Climatol.* **34** 623–42

Harvey C A, Rakotobe Z L, Rao N S, Dave R, Razafimahatratra H, Rabarijohn R H, Rajaofara H and MacKinnon J L 2014 Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar *Phil. Trans. R. Soc.* B **369** 20130089

Hawkins E, Osborne T M, Ho C K and Challinor A J 2013 Calibration and bias correction of climate projections for crop modelling: an idealised case study over Europe *Agric. For. Meteorol.* **170** 19–31

Heft-Neal S, Lobell D B and Burke M 2017 Using remotely sensed temperature to estimate climate response functions *Environ. Res. Lett.* **12** 014013

Holzworth D P *et al* 2014 Apsim—evolution towards a new generation of agricultural systems simulation *Environ. Modelling Softw.* **62** 327–50

Hufkens K, Melaas E K, Mann M L, Foster T, Ceballos F, Robles M and Kramer B 2019 Monitoring crop phenology using a smartphone based near-surface remote sensing approach *Agric. For. Meteorol.* **265** 327–37

Iizumi T, Nishimori M and Yokozawa M 2010 Diagnostics of climate model biases in summer temperature and warm-season insolation for the simulation of regional paddy rice yield in Japan *J. Appl. Meteorol. Climatol.* **49** 574–91

Iizumi T, Okada M and Yokozawza M 2014 A meteorological forcing data set for global crop modeling: development, evaluation, and intercomparison *J. Geophys. Res.: Atmos.* **119** 363–84

Iizumi T, Takikawa H, Hirabayashi Y, Hanasaki N and Nishimori M 2017 Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes *J. Geophys. Res.: Atmos.* **122** 7800–19

Jain M, Singh B, Srivastava A A K, Malik R K, McDonald A J and Lobell D B 2017 Using satellite data to identify the causes of and potential solutions for yield gaps in India's wheat belt *Environ. Res. Lett.* **12** 094011

Li Y, Guan K, Schnitkey G D, DeLucia E and Peng B 2019 Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States *Glob. Change Biol.* **25** 2325–37

Lobell D B and Burke M B 2010 On the use of statistical models to predict crop yield responses to climate change *Agric. For. Meteorol.* **150** 1443–52

Lobell D B, Sibley A and Ivan Ortiz-Monasterio J 2012 Extreme heat effects on wheat senescence in India *Nat. Clim. Change* **2** 186

Mathison C *et al* 2013 Regional projections of north Indian climate for adaptation studies *Sci. Total Environ.* **468–9** S4–17

Menne M J, Durre I, Vose R S, Gleason B E and Houston T G 2012 An overview of the global historical climatology network-daily database *J. Atmos. Ocean. Technol.* **29** 897–910

Mishra V, Shah R and Thrasher B 2014 Soil moisture droughts under the retrospective and projected climate in India *J. Hydrometeorol.* **15** 2267–92

Mondal S, Singh R, Mason E, Huerta-Espino J, Autrique E and Joshi A 2016 Grain yield, adaptation and progress in breeding for early-maturing and heat-tolerant wheat lines in south Asia *Field Crops Res.* **192** 78–85

Morton J F 2007 The impact of climate change on smallholder and subsistence agriculture *Proc. Natl Acad. Sci.* **104** 19680–5

Mourtzinis S, Edreira J I R, Conley S P and Grassini P 2017 From grid to field: assessing quality of gridded weather data for agricultural applications *Eur. J. Agron.* **82** 163–72

Müller C *et al* 2017 Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications *Geosci. Model Dev.* **10** 1403–22

Niles M T and Salerno J D 2018 A cross-country analysis of climate shocks and smallholder food insecurity *PLoS One* **13** e0192928

Oettli P, Sultan B, Baron C and Vrac M 2011 Are regional climate models relevant for crop yield prediction in west Africa? *Environ. Res. Lett.* **6** 014008

Pai D, Sridhar L, Rajeevan M, Sreejith O, Satbhai N and Mukhopadhyay B 2014 Development of a new high spatial resolution (0.25 × 0.25) long period (1901–2010) daily gridded rainfall data set over india and its comparison with existing data sets over the region *Mausam* **65** 1–18

Parkes B, Defrance D, Sultan B, Ciais P and Wang X 2018 Projected changes in crop yield mean and variability over west Africa in a world 1.5 k warmer than the pre-industrial era *Earth Syst. Dyn.* **9** 119–34

Parkes B, Sultan B, Ciais P and Wang X 2017 Modelling fertiliser significance in three major crops *Eur. J. Agron.* **90** 1–11

Porwollik V *et al* 2017 Spatial and temporal uncertainty of crop yield aggregations *Eur. J. Agron.* **88** 10–21

Rajeevan M and Bhate J 2009 A high resolution daily gridded rainfall dataset (1971–2005) for mesoscale meteorological studies *Curr. Sci.* **96** 558–62

Roberts M J, Braun N O, Sinclair T R, Lobell D B and Schlenker W 2017 Comparing and combining process-based crop models and statistical models with some implications for climate change *Environ. Res. Lett.* **12** 095010

Romaguera M, Hoekstra A Y, Su Z, Krol M S and Salama M S 2010 Potential of using remote sensing techniques for global assessment of water footprint of crops *Remote Sens.* **2** 1177–96

Rosenzweig C *et al* 2014 Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison *Proc. Natl Acad. Sci.* **111** 3268–73

Ruane A C, Goldberg R and Chryssanthacopoulos J 2015 Climate forcing datasets for agricultural modeling: merged products for gap-filling and historical climate series estimation *Agric. For. Meteorol.* **200** 233–48

Schlenker W and Roberts M J 2009 Nonlinear temperature effects indicate severe damages to us crop yields under climate change *Proc. Natl Acad. Sci.* **106** 15594–8

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M and Rudolf B 2014 GPCC's new land surface precipitation climatology based on quality-controlled *in situ* data and its role in quantifying the global water cycle *Theor. Appl. Climatol.* **115** 15–40

Semenov M A, Pilkington-Bennett S and Calanca P 2013 Validation of ELPIS 1980–2010 baseline scenarios using the observed European climate assessment data set *Clim. Res.* **57** 1–9

Sheffield J, Goteti G and Wood E F 2006 Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling *J. Clim.* **19** 3088–111

Stackhouse P W, Zhang T, Westberg D, Barnett A J, Bristow T, Macpherson B and Hoell J M 2018 Power release 8 (with gis applications) methodology (data parameters, sources, & validation) *Technical Report* NASA Langley Research Centre

Tack J, Barkley A and Nalley L L 2015 Effect of warming temperatures on us wheat yields *Proc. Natl Acad. Sci.* **112** 6931–6

Tebaldi C and Knutti R 2007 The use of the multi-model ensemble in probabilistic climate projections *Phil. Trans. R. Soc.* A **365** 2053–75

van Oort P, Zhang T, de Vries M, Heinemann A and Meinke H 2011 Correlation between temperature and phenology prediction error in rice (Oryza sativa L.) *Agric. For. Meteorol.* **151** 1545–55

Van Wart J, Grassini P and Cassman K G 2013 Impact of derived global weather data on simulated crop yields *Glob. Change Biol.* **19** 3822–34

Van Wart J, Grassini P, Yang H, Claessens L, Jarvis A and Cassman K G 2015 Creating long-term weather data from thin air for crop simulation modeling *Agric. For. Meteorol.* **209** 49–58

Weedon G P, Balsamo G, Bellouin N, Gomes S, Best M J and Viterbo P 2014 The WFDEI meteorological forcing data set: watch forcing data methodology applied to era-interim reanalysis data *Water Resour. Res.* **50** 7505–14

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N and Kitoh A 2012 Aphrodite: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges *Bull. Am. Meteorol. Soc.* **93** 1401–15

Zaveri E and Lobell D B 2019 The role of irrigation in changing wheat yields and heat sensitivity in India *Nat. Commun.* **10** 1–7