DOI: https://doi.org/10.1016/j.eja.2020.126153

A systematic review of local to regional yield forecasting approaches and frequently used data resources

by Bernhard Schauberger[1], Jonas Jägermeyr[2,3,1], Christoph Gornott[1]*

[1] Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Telegrafenberg A31, 14473 Potsdam, Germany

[2] Department of Computer Science, University of Chicago, Chicago, IL, USA

[3] NASA Goddard Institute for Space Studies, New York, NY, USA

* Correspondence: schauber@pik-potsdam.de; Phone: +49 331 288 20890; ORCID 0000-0001-7917-0392

Co-author mail addresses: jonasjae@pik-potsdam.de; gornott@pik-potsdam.de

Keywords: seasonal crop yield forecasting, literature database, improvement suggestions, weather products, remote sensing, crop mask

Abstract

Forecasting crop yields, or providing an expectation of ex-ante harvest amounts, is highly relevant to the whole agricultural production chain. Farmers can adapt their management, traders or insurers their pricing schemes, suppliers their stocks, logistic companies their routes, national authorities their food balance sheets to guide import or export and, finally, international aid organizations can mobilize reliefs. Evidence has grown in the literature that such forecasts with a meaningful lead time are possible on various geographic scales and for a broad range of crops. Here, we present a systematic review of the methods applied in end-of-season yield forecasting and three frequently used data sources: weather data, satellite data and crop masks. Our literature database comprises 362 studies (2004 to 2019) which were evaluated regarding methods, crops, regions, data sources, lead time and performance. Moreover, we present 24 sources of real-time and predictive weather data, 21 sources of remote sensing data and 16 crop masks. Yield forecasting in our literature sample has been performed for 44 crops in 71 countries, also including many non-staple crops, but with an apparent bias in regions and crops. Forecasting performance depends on various factors, including crop, region, method, lead time to harvest and input diversity. Our systematic review supports a broader application of locally successful approaches at larger scales by providing a comprehensive, accessible compendium of necessary information for yield forecasting. We discuss improvement potentials with respect to methodological approaches and available data sources. We additionally suggest standardization procedures for future forecasting studies and encourage studying additional crops and geographic regions. Implications of forecasts for different target groups on different scales and the adaptation towards climate change are also discussed.

1. Introduction

Estimating several weeks in advance how much there will be on the field at harvest time becomes increasingly tangible. Farmers, commodity markets, insurances, seed traders or logistics companies as well as regional authorities and food aid programs need outlooks on expected harvests to adapt their management of fields, firms or food balances (Basso and Liu, 2019; Ben-Ari et al., 2018; Funk et al., 2019b; Headey, 2011; Johnson, 2014; MacDonald and Hall, 1980; Puma et al., 2015; Stone and Meinke, 2005). An example for the use of forecasts are crop insurances with ex-ante cash transfers or early warning systems, which were ranked as top adaptation measure with the highest economic return on investment (Global Commission on Adaptation, 2019).

There is a multitude of techniques to forecast crop yields, both in the scientific literature and in practical application. Here we present a panoptic of existing and mostly successful approaches from the scientific literature to forecast yields weeks or months ahead of harvest time, covering almost the whole globe and a large range of crops, methods and scales. We also describe available weather data, satellite products and crop masks that facilitate yield forecasting. We chose to include a compilation of weather and remote sensing data due to their importance in forecasting, since there are few studies that do not use at least one of them and also because these are not comprehensively collated elsewhere with a forecasting lens. We furthermore included a compilation of crop masks since these are important for upscaling locally successful forecasting techniques to larger regions, where the risk of confounding due to other crops increases without a proper crop mask. This compilation of results supports upscaling of local approaches, usually on field or region level, to national and global forecasts which are of importance for enabling global food security. It also allows for detecting blind spots of forecasting, i.e. regions and crops which are currently under-represented in the literature but are crucial for global markets and local food security. A comprehensive data base of the selected articles is provided to enable a rapid lookup for other researchers. Finally, drawing of the various approaches presented here, we develop standardization suggestions for robust crop yield forecasting.

Several operational systems on national or regional level for crop yield and production forecasting are in place. These have recently been reviewed by Fritz et al. (2019), Basso and Liu (2019) and van der Velde et al. (2019). They include, among others, the MARS (*Monitoring Agricultural Resources*) system in Europe (van der Velde et al., 2018), Commodity Outlooks by the USDA for the United States and globally (Egelkraut et al., 2015; Johnson, 2014; McKenzie, 2008), CropWatch in China (Wu et al., 2013), ICCYF in Canada (Chipanshi et al., 2015), the Belgian Crop Growth Monitoring System (El Jarroudi et al., 2012; Tychon et al., 2003) or the Indian Mahalanobis National Crop Forecast Centre by the national government (Mahalanobis Centre, 2020).

Global and transnational yield forecasting systems exist, based on models (Iizumi et al., 2018; López-Lozano et al., 2015) or as exchange platforms that combine national forecasts to provide a global outlook (Fritz et al., 2019). Their main aim is to provide a coherent picture of the global harvest situation. The AMIS (*Agricultural Market Information System;* (Bernardi et al., 2016; Delincé, 2017; FAO, 2017)) platform started after the 2007/08 food price spikes to avoid such global wreckage in the future and is supplied, among others, by the GEOGLAM initiative (*Global Agricultural Monitoring*) (GEOGLAM, 2020). Furthermore, there is the FEWS.net platform (*Famine Early Warning System Network*) powered by the US Geo Service to issue early warnings in particular for developing countries (FEWS.NET, 2020; Funk et al., 2019b).

There is only limited literature on the performance of these systems. They often provide accurate forecasts (only a few percent deviation) in average years while having limited capacity to anticipate severe losses in a time frame that would allow for timely intervention (Ben-Ari et al., 2018; van der Velde et al., 2018). Nonetheless, national forecasting systems have improved considerably in recent years, partly owed to increased availability of better data sources on weather, highly resolved remote sensing earth observations, soil, management, and other yield-determining factors, and partly owed to improved methods. Yet the rising frequency of extreme events under climate change (Rahmstorf and Coumou, 2011; Samaniego et al., 2018; Sheffield and Wood, 2008; Stott, 2016) requires a further improvement of techniques.

There is a plethora of methods, mostly on small geographic scale. Apart from a successful small-scale approach, two ingredients are helpful for upscaling these to larger areas: first, the availability of observational weather and/or remote sensing data in near-real time and, second, a high-quality crop-specific crop mask and crop calendars. Not all ingredients are necessary for all approaches, as detailed below (Results section). This study presents a systematic review of the current literature and data availability in these domains. Our aim is to gather and describe existing approaches rather than to evaluate them, as different forecasting purposes (e.g. long lead time, high accuracy, spatial resolution, etc.) would result in different rankings. Our data base is intended to serve as a public repository of approaches and input data that facilitates the application of successful approaches on a larger scale.

This systematic review complements previous reviews on crop yield forecasting. Methods on small geographic scales have been reviewed by, for example, Koirala et al. (2019), Elavarasan et al. (2018), Basso et al. (2013) or Basso and Liu (2019). The latter two provide an overview over different crop modeling schemes, usage of remote sensing data and their applications for yield forecasting. In Basso et al. (2013) the authors distinguish forecasts based on crop observations, statistical or process-based crop models or remote sensing data and list application examples. They also describe possible ways of incorporating remote sensing data into models, and interactions between forecasting and nutrient or pest management. The paper provides a comprehensive overview of early warning systems. Basso and Liu (2019) provide a review of approaches for yield forecasting across the globe, highlighting their virtues and deficiencies with a focus on four staple crops. Their study contains a wide swath of methods and provides suggestions for improved forecasting. Our review, though, goes beyond previous work for the following reasons. First, we perform a fully systematic literature search resulting in a large data base (covering more than 350 articles) including additional (non-staple) crops and allowing only reported data as validation reference. We publish the annotated data base to facilitate further research. Second, we provide reviews of input data, namely operational satellite-borne resources, weather products and crop masks that are helpful for many forecasting efforts. Third, we discuss detailed suggestions on how to improve forecasting approaches in general and specifically for certain crops and regions. This includes suggestions for standardizing future studies on crop yield forecasting to create reliable insights.

The rest of the paper is structured as follows. In section 2, we describe the methods used to compile the four domains of the review (methods, weather data, remote sensing data, crop masks). In section 3, we describe the results obtained for each domain. In section 4, we discuss advantages and drawbacks of our review approach and of the four domains individually. Section 5 concludes.

2. Methods

We performed a systematic review of approaches to forecast yields during the growing season. Systematic reviews are increasingly used in agronomy to present unbiased views of the literature on a subject of interest (Bilotta et al., 2014; Mahon et al., 2017; Makowski et al., 2013; White et al., 2011). Additionally, we compiled comprehensive lists of operational weather reanalysis and forecast data, remote sensing data sources and state-of-the-art crop masks.

In the literature, the word 'forecast' is used in various ways. Newlands et al. (2014) provide a distinction between *forecasting* (a probabilistic statement of future yields with a model), *prediction* (assertion about future yields based on logic), *projection* (future yields based on scenarios) and *prognosis* (subjective judgment of future states). Moreover, we consider *estimation* as a real-time quantification of yield potential, but without forecasting. Throughout this review we use the term 'forecasting' for a pre-harvest statement about crop yields based on one or several models.

In addition to academic efforts, there are operational for-profit solutions for crop yield forecasting. They are not included here, since an exhaustive enumeration is beyond scope and methodological details are often not disclosed.

Apart from the elements reviewed here, historical yields and surveys can aid forecasting. While historical yields, probably with an added trend, may provide early estimates at the beginning of the season, they can usually be replaced by more sophisticated estimates during the season. Surveys among farmers or regional agricultural offices are still a major tool for estimating crop yields throughout the season (Johnson, 2014; Liu and Basso, 2020; van der Velde et al., 2018), but are laborious and difficult to upscale to larger regions. A way out seem picture-based yield estimation techniques (Ceballos et al., 2019), but these have only recently been developed and are therefore not described in this review.

2.1. Systematic literature review of yield forecasting methods

The Web of Science® was queried with the following search terms on various dates until December 03, 2019. As TOPIC constraints we chose "*(crop AND agric*) OR (crop AND yield)*", and as TITLE constraints we selected "*(harvest OR yield OR production OR insur* OR food OR "early warning") AND (forecast OR estimat* OR predict* OR outlook OR pre*harvest OR monitor*)*". Studies were only considered when published between 2004 and 2019. This resulted in 1,334 hits, which were filtered for relevance by subsequently reading title, abstract and full text. The following filter criteria were applied:

- At least one "true" forecast during the season is presented, i.e. before harvest time and without any information on weather or plant growth after the forecasting day.
- Reference yields are observed (experimental study, field data or administrative area) and not simulated, and they are not from greenhouse or pot experiments (except for those crops regularly grown in greenhouses like tomatoes).
- Yield, either absolute or relative, must be forecasted (not biomass or NPP).
- Only food crops are considered.
- The article is written in English, has been peer-reviewed and is identifiable on either the Web of Science® or Google Scholar (the latter was only used to locate the full text of studies indexed, but not linked in Web of Science).
- Review articles were not included, i.e. a new method must be presented or several methods be compared in a quantitative experiment.

These filters resulted in 362 relevant papers. Their contents were tabulated and standardized to enable comparisons and a systematic evaluation by extracting the following variables from each study (not all were applicable to all papers).

- Study data: lead author, year, journal, title
- Experiment setup: study region, crops considered, time frame (harvest years)
- Input data: types of input, sources of input, exogenous variables
- Method: model type, lead time to harvest and spatial resolution
- Evaluation: validation method, performance measures

Regarding method performance, different measures were used in the studies. For standardization purposes, a focus was placed on R^2 (the explained variance) and RMSE (*Root Mean Square Error*), with a split into in-sample and out-of-sample assessments. If none of these measures were available, the study-specific performance was annotated. Several studies mentioned to perform an out-of-sample (OOS) validation, but in their results it was unclear whether these relate to the full data set or OOS; in these cases, the full data set was assumed as training data (i.e. no OOS performance). Some studies, for example Rocha and Dias (2019) or Yang et al. (2019), use OOS to tune model hyper parameters, which is not counted as OOS forecast if for the latter the full data set was used.

2.2. Collection of sources of weather data

Yield forecasting based on crop growth models, whether statistical or process-based, usually requires weather data as input. Based on the considered literature and our own knowledge, we compiled an overview of commonly used near-real time products documenting past weather conditions until the forecasting date as well as sources of weather forecasts up to several months ahead. Our compilation of weather products comprises global and major regional efforts.

2.3. Collection of sources of remote sensing data

Data from remote sensing tools are used in many forecasting studies. These tools vary in observed bandwidths, spatial resolution, temporal revisit frequency, distance to the measured object and, not least, cost for acquisition. Hand-held sensors, tractor-mounted devices or drones are used in several local forecasting studies, which we also classified as remote sensing within this study. The major source of data for larger areas at reasonable cost are, however, satellites. To illustrate the wide range of available satellites for crop monitoring, a list of these was compiled. It was populated with satellites referenced in the studies selected above and further satellites described on overview sites of NASA, ESA, Chinese and Indian space agencies.

2.4. Collection of crop masks

Crop-specific growing areas are required to derive crop production over a region. Masks are not a direct input to forecasting methods, but are used to filter out irrelevant areas where the crop of interest is not grown, either in pre- or post-processing. There are several national and global products available, which were collected and described according to their resolution, update schedule, crop specificity, method for construction and resources used. Their quality is not evaluated here as the applicability of a certain mask may depend on the usage. The list of crop masks was compiled from both the article library compiled above and additional review articles on crop masks (see references in the section). Only crop masks with global or at least multi-national coverage are listed, with an exception for the USA where many products are available and crop production is important globally.

3. Results

3.1. Methods for yield forecasting used in the literature

The selected articles on yield forecasting show a wide range of crops (n = 44), including non-staple or horticultural crops like citrus or strawberry, but with less coverage (Figure 1). Wheat, maize and rice are among the most studied crops.

Studies have been conducted in 71 countries (Figure 2), most frequently in the USA, followed by China, India, Spain and Brazil. There are also many experiments in developing countries, but often only with a single study on a single crop. Forecasting efforts in Europe are dispersed across space, largely following country size and production share, with a dearth of studies particularly in Eastern Europe. An overview of all combinations of crops and study countries is provided in Figure S1.

The number of years considered per study varies between 1 and 147 (average 10.6) and depends on the study type: regional studies based on aggregated yield data are usually longer than experiment-based assessments (Figure S2).

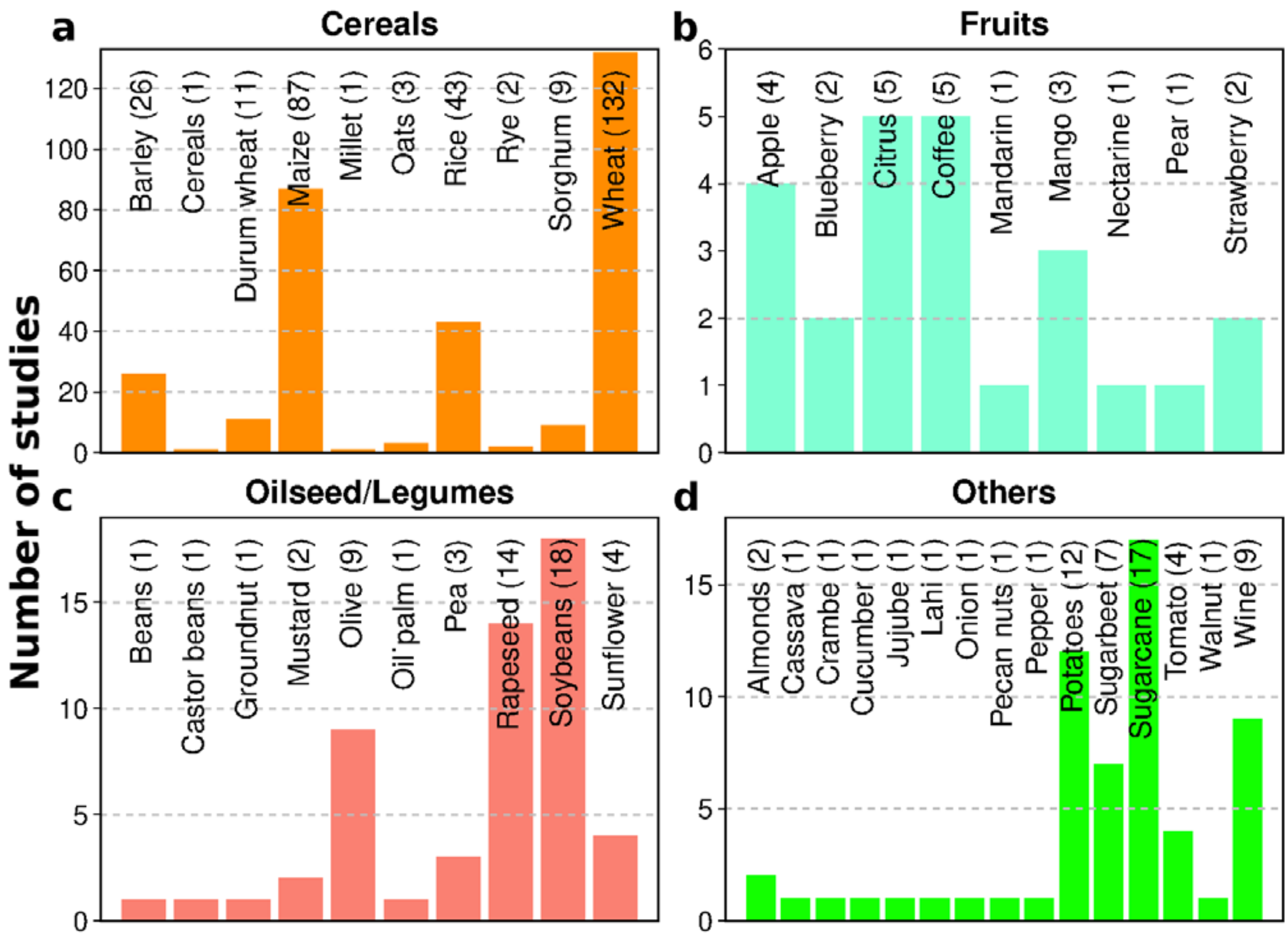

Figure 1: Crops for which yield forecasts were performed. The y axis denotes the number of studies for each crop (also indicated in parentheses behind the crop name). Several crops may be addressed by the same study, so the total sum is larger than the number of reviewed studies. Y-axis ranges differ between categories.

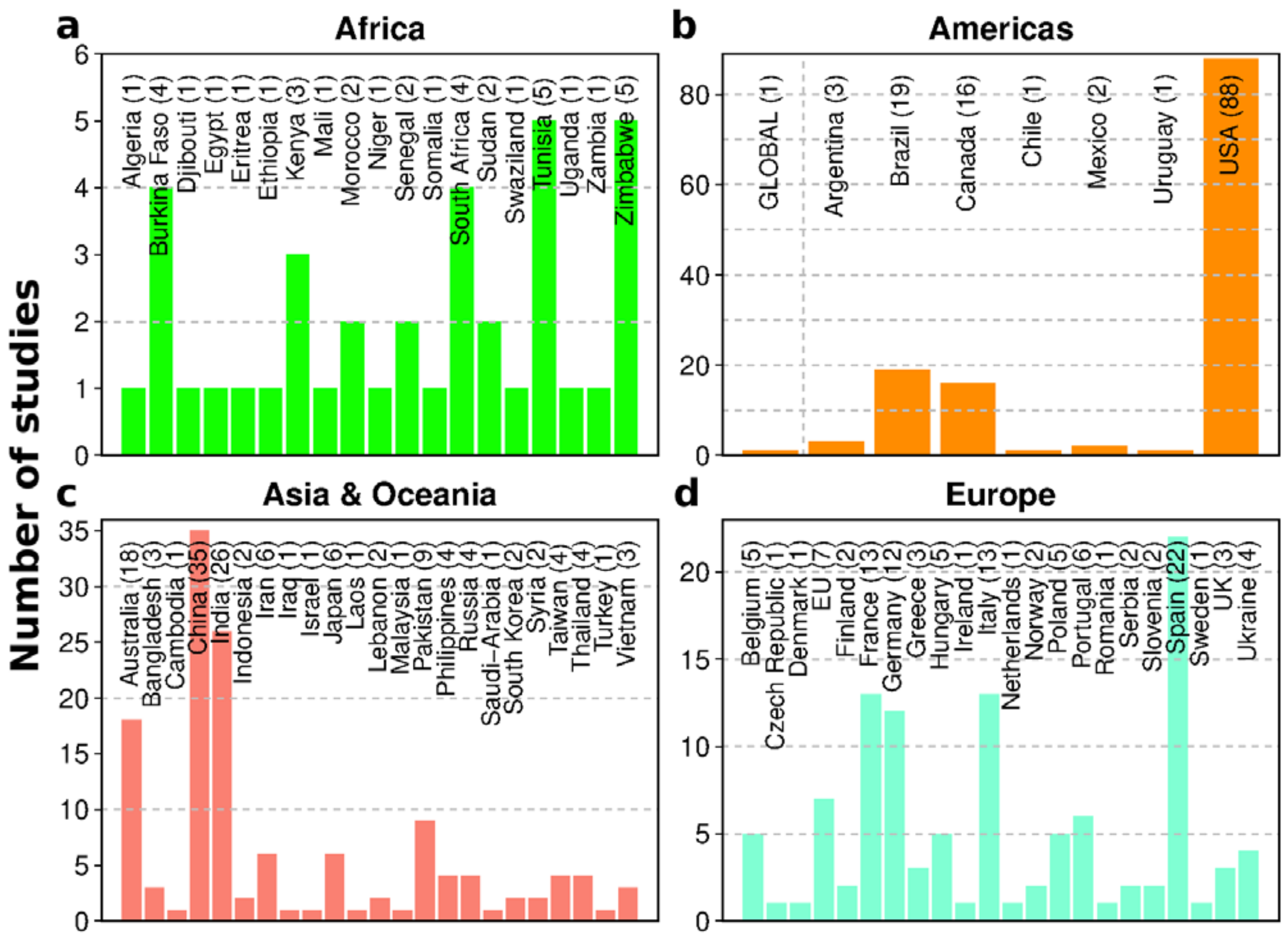

Figure 2: Countries in which yield forecasting was conducted. The y axis denotes the number of studies addressing a country (also indicated in parentheses behind each country); several countries may be addressed by the same study. A majority of articles only studies sub-regions within one country. There is one global study that we included in panel b. Y-axis ranges differ between regions.

There is a wide range of model types and specific models used for forecasting (Figure 3), with a bias in favor of regression models (occurring in 74% of studies). More than one method is applied in 37% of the studies (Figure S3). Process-based crop models occur in 57 studies in our data base, among which DSSAT (including CERES) with 13 mentions is most prominent. Delincé (2017) and Basso and Liu (2019) provide an overview over the use of crop models in forecasting. Within the machine-learning models, neural networks are featured most.

A variety of data enters the models as exogenous variables or as calibration targets for intermediate outputs (Figure 4). If weather data are used as model input, these are often combinations of temperature, precipitation, radiation, or humidity (herein coined as ‘Multiweather’ for standardization). Similarly for vegetation indices (VI), a number of studies uses multiple VIs and tests which one fits the yield data best (termed ‘MultiVI’). The NDVI (Normalized Difference Vegetation Index) is the most prominent single VI used for yield forecasting. Within direct measurements, LAI (Leaf Area Index) is most often used. Various combinations of different inputs, for example remote sensing and weather data, are used (Figure S3). Remote sensing data alone (125 times) and weather data alone (75) are most often used, followed by their combination (55). A split of input combinations per model category highlights the differential usage between, for example, process-based and statistical models (Figure S4).

Scientific efforts in yield forecasting have increased (Figure S5). Model usage has evolved between 2004 and 2019, with rising frequency of machine-learning methods like Neural Networks, SVM or Random Forests (Figure S6).

Most studies were performed on regional scale (163 studies, where ‘regional’ indicates any administrative unit below country level), followed by 127 experimental plot studies and 61 field-scale studies. Thirteen studies were performed on national, one on global and one on storage silo level. Four studies treated more than one scale.

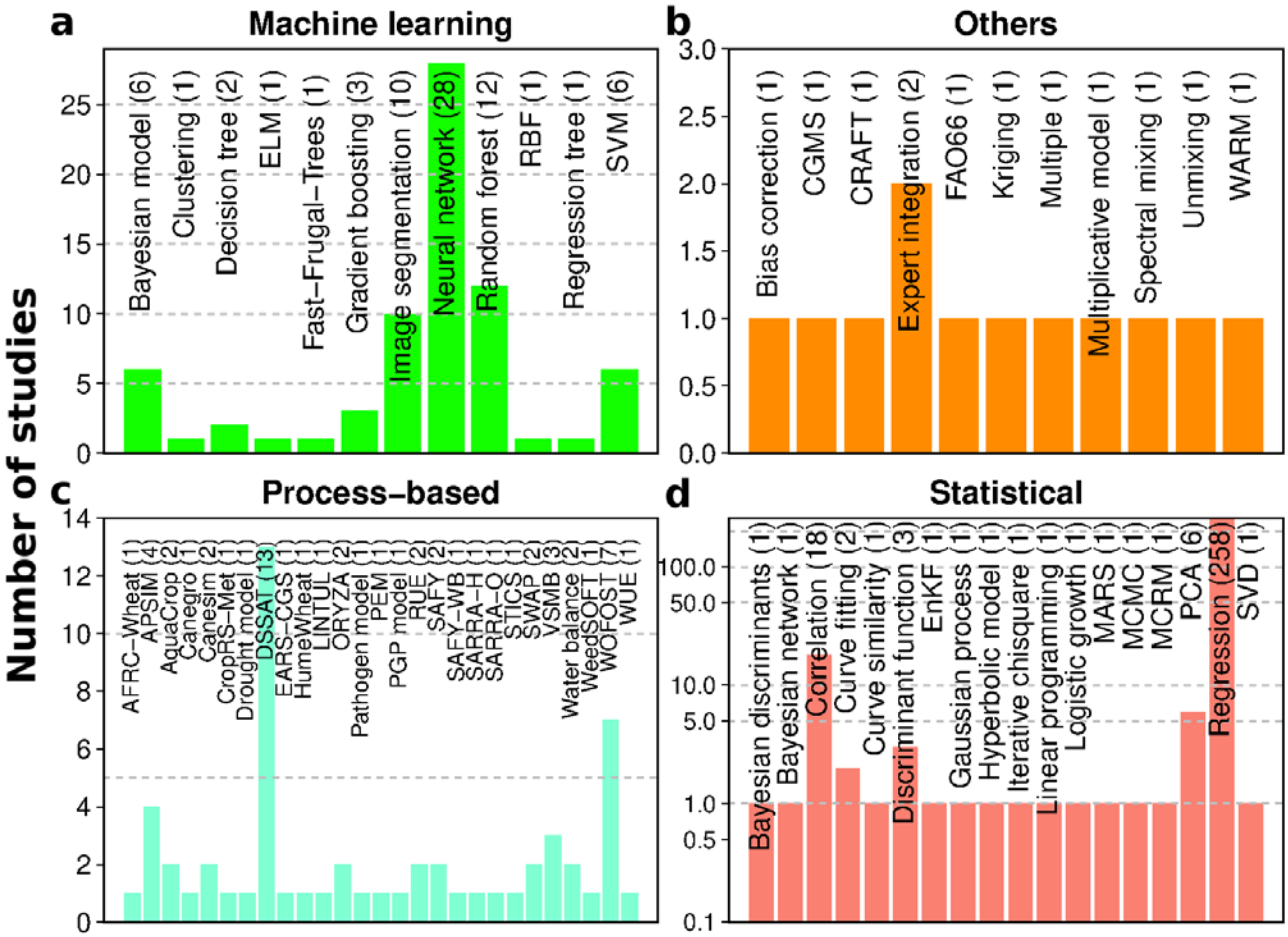

Figure 3: Overview of the methodological tools in the studies considered. Methods are grouped in four broad categories; "Others" are methods that do not fit elsewhere. The y axis denotes the number of times a technique or model is used in the literature (one study may use several tools). Y-axis ranges differ between categories; the y-axis for panel d is log-scaled to accommodate the wide range.

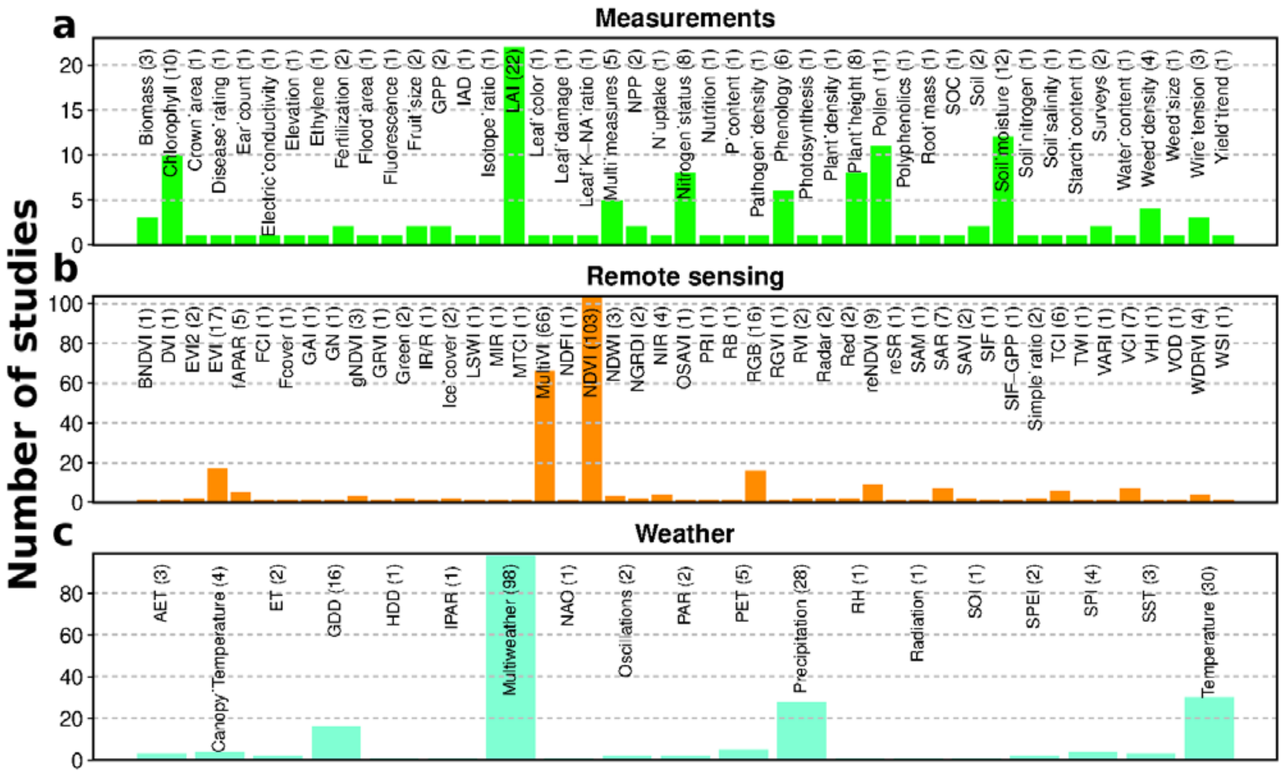

Figure 4: Exogenous variables used as model input, grouped intro three broad categories. The y axis denotes the number of times an input is used in the literature (one study may use several tools). Y-axis ranges differ between categories. A list of abbreviations is provided in Table S1.

Lead times to harvest range from just before harvest up to one year or before season start. The exactness of lead times indicated in the studies varies, with 26 studies mentioning only that a forecast before harvest was performed, but the actual lead time is not stated. The most frequent lead times are two months (71 articles) and one month (70). Forecasts are provided at various growth stages, with the longest lead times (six months or more) observed for olives, sugarcane, coffee and citrus.

The forecasting performance depends on the crop, region, lead time to harvest, inputs used and method applied (Figure 5). Due to the large variety in reporting performance, we focused on those 134 studies which clearly indicated the time point of forecasting, reported in-sample R^2 values per crop (out-of-sample values were too infrequent) and considered only six crops with high global production share.

The majority of studies (52%) does not provide an out-of-sample assessment of their performance. Out-of-sample means that a forecast method was trained on a subset of the data and then applied without further calibration to a test set, as would be the situation for an operational forecast. Furthermore, several studies that indicate to have done an out-of-sample validation do not state the performance of this exercise. Of those studies claiming to have done a validation with independent test data (173), only 109 (30%) report these performance measures as R^2, RMSE or similar.

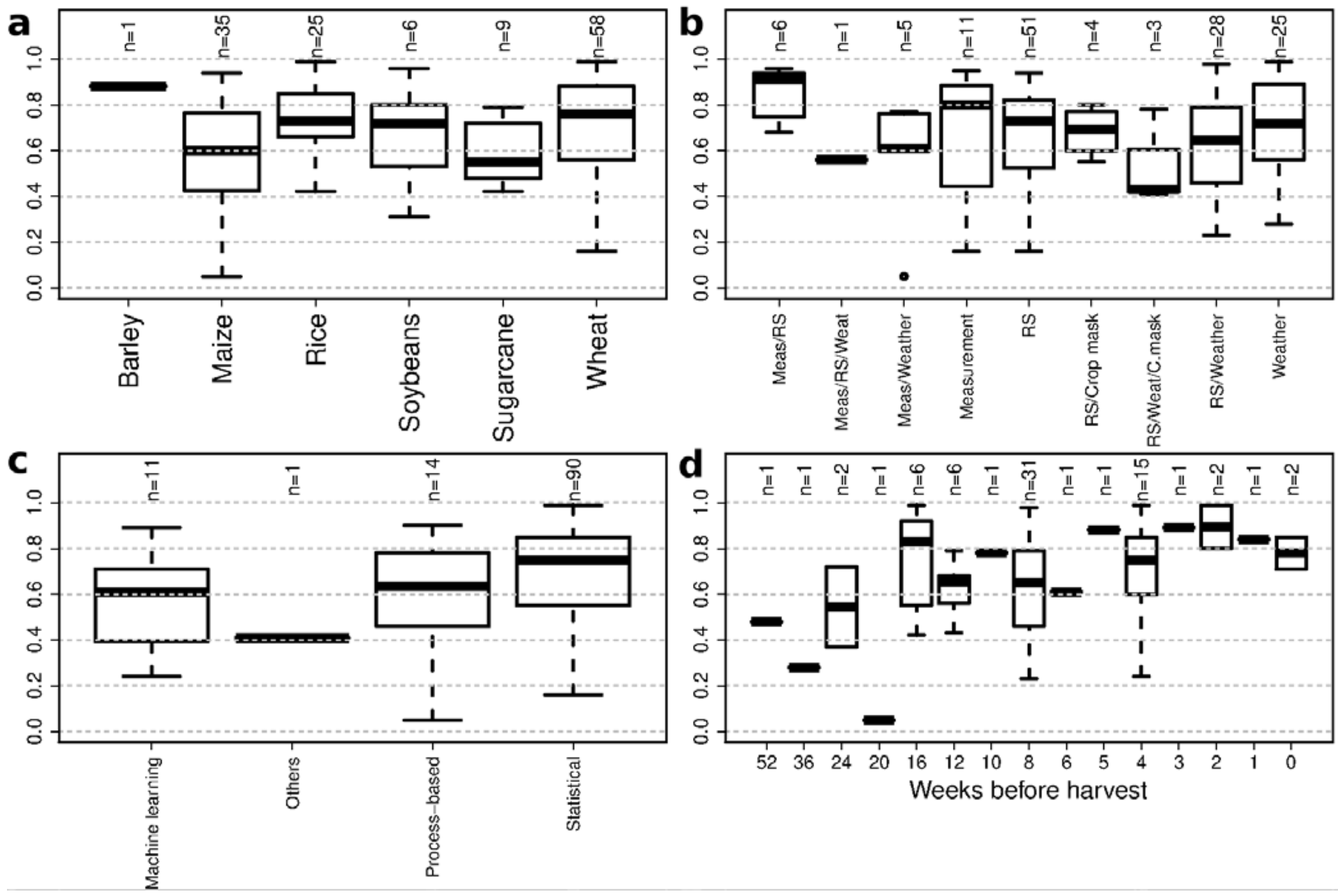

Figure 5: Dependency of forecast performance (R^2 in-sample; y axis) on different determinants: (a) crops, (b) input types, (c) method types and (d) lead time to harvest. For the

methods plot (c) only those entries were considered where the R^2 *could clearly be associated with one method. For the lead time plot (d) only those entries were considered where lead time could be normalized to weeks before harvest. Abbreviations in (b) are C.mask = Crop mask, Meas = Measurement, RS = Remote Sensing, Weat = Weather*

3.2. Sources of weather data

In Table 1 we provide a list of the most common near-real time observation-based weather products, plus an overview over short-term and seasonal weather forecasts. NASA POWER (Prediction Of Worldwide Energy Resources) is a prominent near-real time product with a latency of few days for most variables, combining solar radiation data based on radiative transfer models from satellite observations with meteorological data from MERRA2 (Modern-Era Retrospective analysis for Research and Applications, version 2). The time gap (latency of few months) until MERRA 2 becomes available is bridged with GEOS FP (Global Earth Observing System Forward-Processing) to provide near-real data. ERA5 is the new ECMWF (European Centre for Medium-Range Weather Forecasts) near-real time reanalysis product, replacing ERA-Interim reanalysis (discontinued on 31 August 2019). CHIRPS (Climate Hazards Infrared Precipitation with Stations) and CHIRTS (Climate Hazards Infrared Temperature with Stations) are new high-resolution products at 3 arcmin global coverage, combining satellite observations with station data. These products provide an advantage especially in regions with scarce station data.

In addition to global products there are various regional efforts. In Table 1 we list a selection of the prominent North American and European products, but this list is not meant to be exhaustive. There are, for example, near real-time data sets by the Indian Meteorology Department (IMD). We also refrain from collating static, not regularly updated reanalysis products and climate forcing data sets, which are discussed in Ruane et al. (2015).

Several options are available to continue yield simulations after the forecasting day throughout the growing season (Basso and Liu, 2019). These include historical weather data, employing a weather generator, or using weather predictions. For the latter category we provide an overview of the most common weather forecast products currently available. General short-term weather forecasts (up to 2 weeks) can be continued with sub-seasonal (up to 3 months) and seasonal outlooks, but forecast skill beyond 10 days is as yet generally marginal (Bauer et al., 2015; Kushnir et al., 2019). Prominent short-term products include NCEP's GFS (abbreviations in Table 1) in the U.S. and the ECMWF ensemble. The TIGGE project assimilates data from ten global numerical weather prediction centers. The field of sub-seasonal to seasonal predictions is under active development, complementing operational products hosted by NCEP and ECMWF with new efforts such as the S2S project across multiple institutions (references in Table 1).

Several studies in our literature data base use historical climate as a proxy for unobserved weather in the future. This can be achieved via a trend extrapolation, historical averages or more sophisticated choices where the observed weather during the season of interest until the forecasting day is used to identify historical weather analogues (i.e. similar weather until that day) to prescribe a weather trajectory until the end of the season (Anwar et al., 2008). Teleconnections with, for example, sea surface temperatures are an emerging option for forecasts (Boers et al., 2019; Knippertz et al., 2003; Lehmann et al., 2020).

Table 1: Overview of weather products useful for yield forecasting. The table is split into three sections: global near real-time observational products, regional near-real time products and (mostly) global short- to long-term forecasts. Variable codes are: 2m air temperature (T), precipitation (P), humidity (H), shortwave (SW) and longwave (LW) surface radiation, wind (W).

Name	Time period	Latency	Temporal resolution	Spatial resolution and coverage	Variables	Comments	Ref.
Global near real-time products							
NASA POWER (Prediction Of Worldwide Energy Resources)	1981 to present	1-7 days	Daily (initially 1-hourly (meteorological) and 3-hourly (solar) but averaged to daily)	0.5° global (initially 0.5° x 0.66° (meteorological) and 1° (solar), then rescaled to 0.5°)	T, P, H, SW, LW, W	Solar data based on radiative transfer models from satellite observations, using NASA GEWEX SRB and FLASHFlux, meteorological data taken from MERRA 2 and GEOS 5.12.4	(NASA POWER, 2020)
NASA POWER - **GEWEX SRB 3.0** (Global Energy and Water cycle Exchanges Surface Radiation Budget) and NASA POWER – **CERES FLASHFlux 2 & 3** (Clouds and the Earth's Radiant Energy Systems Fast Longwave and SHortwave Radiative Fluxes)	1983 to 2007 and 2008 to present	7-8 days	3-hourly	1° global	SW, LW	Satellite based, filling the gap until highly accurate CERES flux estimates become available 6-12 months later	GEWEX (2020)
NASA POWER - **MERRA 2** (Modern-Era Retrospective analysis for Research and Applications, version 2)	1981 to present	Few months	hourly	0.5° x 0.625° global	T, P, SW, LW	NASA's Global Modeling and Assimilation Office (GMAO) reanalysis product, in-situ and satellite data observations, replaces	Gelaro et al. (2017)

Name	Time period	Latency	Temporal resolution	Spatial resolution and coverage	Variables	Comments	Ref.
						MERRA and now provides microwave, hyperspectral, and ozone outputs	
NASA POWER - **GEOS FP** (Global Earth Observing System Forward-Processing)	End of MERRA 2 to near-real time	About 4 hours	hourly	0.25° x 0.3125° global	T, P, SW, LW	NASA's Global Modeling and Assimilation Office (GMAO) operational product, used to fill gaps between end of MERRA 2 and near-real time, kept available for approximately 6 months	Borovikov et al. (2017)
ERA5 (ECMWF ReAnalysis)	1950 to present, including a forecast field	3 months, preliminary data 7 days	hourly	0.25° global	T, P, SW, LW, W	Observations and modeling, replaces the ERA-Interim reanalysis stopped in 2019, ERA5 will be completed by mid 2020	ERA5 (2020)
NCEP/NCAR (National Center for Atmospheric Research) Reanalysis 1	1948 to present	1 day	Daily, 6 hourly	2.5° global	T, P, H, W	Climate Data Assimilation System (CDAS)	Kalnay (1996)
NASA GLDAS (Global Land Data Assimilation System)	1948 to present	About a month	3-hourly	0.25° global (all land north of 60° S)	T, P, SW, LW	Assimilation of satellite- and ground-based data into land-surface models (uncoupled from an atmospheric model) forced with observations, and thus not affected by numerical weather prediction forcing biases.	Rodell et al. (2004)
CHIRPS (Climate Hazards Infrared Precipitation with Stations)	1981 to present	preliminary 2 days, final 3 months	daily	0.05° (3 arcmin) global	P	Satellite and station data for high-resolution estimates in regions scarce observations	Funk et al. (2015)

Name	Time period	Latency	Temporal resolution	Spatial resolution and coverage	Variables	Comments	Ref.
CHIRTS (Climate Hazards Center Infrared Temperature with Stations)	1983 to 2016 (to present soon)	Not yet operational	daily	0.05° (3 arcmin) global	Tmin, Tmax	geostationary satellite thermal infrared and 15000 stations, high-resolution Tmax, advantage for near-real time	Funk et al. (2019a)
NASA GPM-IMERG (Integrated Multi-satellitE Retrievals for Global Precipitation Measurement)	2000 to present	6 hours	30 min	0.1° global	P	Integrated multi-satellite retrievals of precipitation and snow coverage. TRMM (Tropical Rainfall Measuring Mission) and CHIM was discontinued in 2015 and replaced by GPM, it included Multi-satellite Precipitation Analysis (TMPA), the real-time TMPA (TMPA-RT)	Huffman et al. (2019)
NCEP – CFSR *(National Centers for Environmental Prediction - Climate Forecast System Reanalysis)*	*1979 to 2017*	*discontinued*	*6 hourly*	*0.5° global*	*T, P, SW, LW*	*Coupled atmosphere-ocean-land surface-sea ice system, has been extended as an operational real-time product (CFSv2)*	*Saha (2010)*
Regional near real-time products (selection)							
NASA NLDAS (North American Land Data Assimilation System)	1979 to present	About 3-4 days	Daily, hourly	7.5 arcmin (1/8°), North America	SW, H, W	Part of GLDAS, at higher spatial resolution for North America	NLDAS (2020)
ORNL DAAC DayMet (Oak Ridge National Laboratory Distributed Active Archive Center)	1980 to present	3-4 months	Daily	1 km, North America, Hawaii, Puerto Rico	T, P, SW, H	Gridded estimates of daily weather parameters	DayMet (2020)
PRISM (Parameter-elevation	1981 to	Released	Daily, monthly	30 arcsec, USA	T, P	Assimilated weather station data,	Daly et al.

Name	Time period	Latency	Temporal resolution	Spatial resolution and coverage	Variables	Comments	Ref.
Regressions on Independent Slopes Model)	present daily values, 1895-present monthly values	1pm EDT for preceding day				data released in "early release" which can change, 7 months are considered "provisional release", then "stable"	(2008)
Copernicus E-OBS	1951 to present	About a month	Daily	0.1° Europe extended	T, P, radiation	European high-resolution ensemble dataset	Cornes et al. (2018)
Global short-term to seasonal forecasts							
NCEP GFS (Global Forecast System)	8 days at 0.5° and up to 16 days at 1°	6 hours	3-hourly	0.5° global	T, P, H, W	U.S. numerical weather prediction system, one of the predominant synoptic scale medium-range models	GFS (2020)
TIGGE (The International Grand Global Ensemble)	1 to 15 days	48 hours	6-hourly	Global	T, P, SW	Previously named THORPEX, ensemble forecast data from 10 global numerical weather prediction centers, including NCEP, ECMWF, UKMO, etc.	TIGGE (2020)
ECMWF – HRES (High Resolution) and **ECMWF – ENS** (Ensemble)	10 and 15 days		Twice daily	9 and 18 km global	T, P, H, SW, W	European numerical weather prediction system, ensemble of 51 forecasts	ECMWF (2020)
ECMWF – Extended and Long Range	46 days and 7 months		Twice weekly and monthly	36 km global	T, P, H, SW, W	Extended range: bridges gap between medium-range and seasonal forecasting, but still with land-atmosphere-ocean coupling	[as above]

Name	Time period	Latency	Temporal resolution	Spatial resolution and coverage	Variables	Comments	Ref.
						Long range: ensemble of 51 forecasts available on the 8th of each month, in addition to monthly product, run quarterly for 1 year ahead	
S2S (Sub-seasonal to Seasonal Prediction Project)	Up to 60 days	Updated daily with a 21 day delay	Weekly anomalies	Global	T, P	Ongoing joint research project across multiple institutions within WWRP/THORPEX/ WCRP	Vitart and Robertson (2018); S2S (2020)
NCEP - CFSv2 (Coupled Forecast System model version 2)	45 days to 6 months	6 hours	6-hourly	0.5° global	T, P, SW, LW,	Medium to long range numerical weather prediction and a climate model to bridge weather and climate timescales, provides anomalies with respect to 1999-2010 climatology;, "Coupled" refers to the fact that the model couples atmosphere and ocean	Saha et al. (2014)
CHFP (The Climate-system Historical Forecast Project)	sub-seasonal-to-decadal	Under development	Under development	Global	Under development	Anomalies, multi-modal and multi-institutional experimental framework for sub-seasonal-to-decadal complete physical climate system prediction, including atmosphere, oceans, land surface and cryosphere	Tompkins and al (2017)

Name	Time period	Latency	Temporal resolution	Spatial resolution and coverage	Variables	Comments	Ref.
NOAA CPC (Climate Prediction Center)	2 weeks to 13 months	Few days	Monthly	conterminous United States	T, P	Maps of anomalies with respect to 1980-2010 average climatology	NOAA CPC (2020)
JMA/MRI-CPS2 (Tokyo Climate Center, Japan)	Up to 7 months	Few days	Daily, later monthly	0.5° global	T, P, W, H	ENSO forecasts, seasonal anomalies; hindcast archive 1979-2014	Takaya et al. (2016)

3.3. Sources of remote sensing data

A majority (66%) of the studies reviewed here uses remotely sensed data to produce yield forecasts during the season. While 104 small-scale studies on experimental plots relied on ground-based (74) or low-height airborne (30) sensors, the majority used data from satellites. Only a few satellites account for most applications (Table 2): MODIS (67 studies), SPOT (27), Landsat (26) and AVHRR (21). There are more satellite products suitable for application, but used only infrequently or not at all. A selection of these is provided in Table 2. Stratoulias et al. (2017) and Chivasa et al. (2017) provide further overviews of satellite sensors with very high spatial resolution and how to estimate yields in heterogeneous landscapes. Depending on the wavelengths measured, infrared, optical, radar, microwave and multi-/hyperspectral sensors are distinguished.

Applications of these data within yield forecasting comprise the measurement of vegetation greenness, general vegetation condition, soil attributes like soil moisture, pest or disease infestations (often based on greenness) or crop distribution. Weather data can also be collected with satellites, for example METEOSAT.

Remotely sensed data can be used, as training or validation inputs, in three different ways: direct (for correlation), indirect (via assimilation into models) or as plant measurement surrogates (for further processing in models); see also Basso and Liu (2019).

A plethora of different vegetation indices or conditions can be derived from satellite measurement, of which many were used in our database (Figure 4). The NDVI accounts for the large majority, and 66 studies use a combination of at least three spectral measurements to forecast yields. Cao et al. (2015) and Basso and Liu (2019) provide an overview of many possible indices. A recent effort to harmonize Landsat and Sentinel-2 data (HLS, 2020) has been started by NASA to overcome data gaps between both products.

Table 2: Satellite sources for remote sensing data with key characteristics and their usage count within the studies reviewed here; abbreviations are listed in Table S1. URL resources for each satellite family are provided in Table S4. Application references for each satellite used within the studies reviewed here are detailed in Table S2 (not all satellites were applied in our literature set).

Satellite family	Operator	Spatial resolution	Revisit frequency	Spectral bands	Area coverage	Availability	Operational time frame	Studies with use
AVHRR	NOAA	1 km	1 day	NDVI	Global	Public	1981-present	21
CBERS	China-Brazil Earth Resource Satellite Program	Up to 10 m	26 days	B,G,R, NIR	Global	On request	1999-present	0
COSMO	Italian Space Agency	Up to 1 m	16 days	SAR	Global	On request (public)	2008-present	0
Deimos	Deimos Imaging	22 m	2-3 days	G, R, NIR	Global	On request (comm.)	2009-present	0
ENVISAT	ESA	Up to 30 m	35 days	SAR C-band	Global	On request	2002-2012	4
Formosat	NSPO (Taiwan Space Agency)	8 m	1-2 days	B,G,R, NIR	Global	On request (comm.)	2004-present (gap in 2016/17)	3
IKONOS	DigitalGlobe	1-4 m	3-14 days	B,R,G, NIR		On request (comm.)	1999-2015	0
IRS	ISRO (Indian Space Research Organisation)	6-70 m	24 days	R, G, NIR, SWIR	75 N, 25 S, 20 W, 50 E	On request	1996-2013	2
KOMPSAT	KARI (Korea Aerospace Research Institute)	4-6 m	14 days	B,R,G, NIR	Global	On request	1999-present	0
Landsat	USGS, NASA	30 m	16 days	B,G,R, NIR, SWIR, TIRS	Global	Public	1982-present	26
MODIS	NASA	Up to 250 m	1-2 days	36 bands	Global	Public	1999-present	67
Pleiades	Airbus	2 m	Up to daily	B,R,G, NIR	Global	On request (comm.)	2011-present	0
Proba-V	ESA	0.3-1 km	1-2 days	NDVI	56S-75N	Public	2013-present	1
Quickbird	DigitalGlobe	2.5 m	Up to 2.5 days	B,G,R, NIR		On request (comm.)	2001-2015	2
Radarsat	Public-private	1-100 m	Daily	SAR C-band	Global	On request	1993-present	6

Satellite family	Operator	Spatial resolution	Revisit frequency	Spectral bands	Area coverage	Availability	Operational time frame	Studies with use
	cooperation from Canada							
RapidEye	Planet	6.5 m	Up to daily	B,G,R, RE, NIR	84 N to 84 S	On request (comm.); partly public	2009-present	4
ResourceSat-1/2	ISRO (Indian Space Research Organisation)	6-70 m	5 days	G, R, NIR, SWIR	Global	On request	2004-present	0
Sentinel	ESA	Up to 5 m (SAR) or 10 m (other)	5-12 days	SAR, 13 bands	Global	Public	2014/5-present	5
SPOT	Spot Image / CNES	6-20 m	Up to daily	B,G,R, NIR	Global	On request (comm.)	1986-present	27
TerraSAR, TanDEM	German Aerospace Center (DLR)	< 1 m to 40 m	11 days	SAR	Global	Public	2007-present	2
WorldView	DigitalGlobe	< 1 m	Depends on region, a few days	B, R, G, NIR, +4 more	Global	On request (comm.)	2007-present	2

3.4. Crop masks of potential relevance for yield and production forecasting

A crop mask is a gridded product, indicating the location or grid-cell fraction of a specific crop. Crop locations are important for selecting areas of interest in yield or production forecasting and calculating aggregated production from yield. These masks are usually applied in pre- or post-processing of forecasting methods, not as a direct input to the forecasting algorithm. There are several global and regional products publicly available, offering different spatial resolutions, land-use classifications, update schedules and resources used for construction (Table S3). Most of the masks are static (created for one specific year) and distinguish only land-use classes like forests or croplands, but do not resolve different crop types in higher granularity. A majority of masks provides geo-specific locations for land-use classes, indicating the dominant land-use class for each pixel. Few masks, like SPAM or MIRCA2000, indicate the area shares of each crop type per grid cell but do not spatially allocate these types within grid cells. Some of the masks, like the MODIS-LC or FAO's WaPOR products, offer further data like phenology or evapotranspiration with regular updates. MIRCA2000 also provides a crop calendar including monthly planting and harvest dates. Almost all masks are constructed with the support of remote sensing data, either as sole resource or to upscale ground-based data to a larger area. Comparisons of different approaches and masks are provided by Grekousis et al. (2015), Anderson et al. (2015), Lambert et al. (2016) and Fritz et al. (2010). Apart from the crop masks listed here, there are many national approaches, exemplified by Pervez and Brown (2010) or Waldner et al. (2017). Within our literature data base, the US-based CDL was most often used (15 studies), followed by MODIS Land cover (5). Others were used only once or not mentioned.

For deriving production forecasts for the ongoing season, an updated crop mask during the season is necessary. This is not the case for any of the masks described in Table S3, as these are released only after the season. This leaves two choices for obtaining a current mask: either assume the latest available data of crop allocation as still valid, or use a dynamic approach that allows for an on-the-fly construction of a crop mask, based on ground observations, the methods used for constructing the existing masks or idiosyncratic algorithms. The Sen2Agri toolbox (Valero et al., 2016), neural network-based crop mapping in California (Zhong et al., 2019), a global comparison of five methods for satellite-derived crop masks (Inglada et al., 2015) or random forest-based crop masking in Zambia with the Google Earth Engine (Azzari and Lobell, 2017), among others, are of interest here. Another approach is yield correlation masking (Kastens et al., 2005), which filters crop areas by correlating reported crop yields with vegetation indices for all potential areas and keeping only those with a minimum correlation.

While the highest resolved masks on global level (30 m) are sufficient for large fields in, for example, the USA, Australia or Ukraine, they are usually not sufficient for many developing countries in particular in sub-Saharan Africa where plot sizes are small and often under mixed or inter-cropping regimes with two crops grown simultaneously or in relay. If no data with higher resolution are available, an un-mixing approach might be useful to distinguish the contributions of individual crops to a larger, mixed observed pixel in the mask (Rembold et al., 2013). In the tropics, in-season updates of crop masks are additionally hampered by large parts of crop growth being in the rainy season. Clouds block the penetration of non-radar wavelengths and thus preclude eliciting crop type, phenology or growth. Diverse and complex cropping patterns in the tropics lead to patchy influences from weather and management. Additionally, less dense data coverage for yields, management or other required information (Kamali et al., 2018) complicates model calibration.

4. Discussion

We presented a systematic review of crop yield forecasting methods and three often used data domains: weather, remote sensing and crop masks. This compilation can aid in upscaling successful local approaches to larger regions and thus support decision making in agriculture on multiple levels.

4.1. Advantages and concerns of our systematic literature overview

The search terms for the Web of Science® were chosen to gather a broad scope of studies related to forecasting. Yet the selected articles may be biased with respect to crops, regions, performance and/or inputs considered. Additionally, for practical reasons but also to synthesize recent developments we limited the selection of studies to a publication date between 2004 and 2019, which excludes important previous studies such as Cane et al. (1994). For horticultural crops like strawberries or tomatoes only few studies were found, which might be explained by the '*agric**' search term. Biotic stressors such as pests and diseases or the detection of phenological phases are rarely addressed in the studies reviewed, due to our focus on yields. This merits an additional study to catalogue recent efforts on that front. Survey-based methods or expert judgments are only sparsely included in our data base – which supports our aim of compiling a suite of methods that are technically ready for upscaling. The literature in general may be selective in terms of successful performance, as methods with low performance are unlikely to be published – though these could hone expectations for certain crops and regions. Moreover, we excluded all studies that estimate yields after harvest to focus on true forecasts during the season (or just before harvest time), although real-time estimations shortly after harvest could improve regional statistics which often appear with a significant time lag. Several studies were excluded as forecasts were not compared with observed but rather simulated yields (using full-season data), which we assumed as biased towards better performance as both forecast and full-season simulation tend to capture similar signals (examples: Ferrise et al. (2015), Brown et al. (2018)). Nonetheless, for method improvement or data assimilation techniques such studies could provide relevant insights. Finally, it was difficult to standardize a very diverse range of methods, performance assessments or input data. For example, there are many flavours of regression analyses (OLS, PCR, PLS, stepwise etc.) which were not specifically annotated.

We assume that the 362 studies reviewed represent the current state of the art in crop yield forecasting techniques. Moreover, as no filter on crops or methods was imposed, this review gathers a broad scope of published techniques for very diverse crops. There is a partial overlap between our article data base and the one by Basso and Liu (2019). Differences are owed to the search method, terms, time frame and filters, which renders both studies complementary.

With the publicly available result table (Table S2), this review may help to quickly select appropriate methods or data sources and standardize efforts for future yield forecasting approaches for specific crops and regions.

4.2. Strengths and weaknesses of current yield forecasting methods

There are several studies that reach a very high accuracy of forecasting, i.e. more than 90% explained variance or an RMSE of less than 5%. These methods comprise different approaches and input data. Many successful studies are placed in water-limited environments where a rainfall deficit directly affects yields (Balaghi et al., 2008; Guo and Xue, 2011; López-Lozano et al., 2015; Maresma et al., 2016).

Robust correlations between forecasted and actual yields are reported for methods based on only remote sensing data or only weather data and both combined. Empirical correlations between remote sensing-derived data and yields are comparatively easy to construct and often show robust performance, across many different regions. Regression approaches have the advantage that they usually do not require soil or cultivar parameters. Yet, detailed information on growing season and phenology (possibly derived from satellites) often improves the model. For models based only on weather, the necessary input data can partly be replaced by satellite-measured weather data, if ground-based information is not available (an example is the CHIRPS data set, Funk et al. (2015)). Combined models, using both remote sensing and weather data, unite the advantages of each, but may suffer from collinearity between exogenous inputs. Shorter lead times to harvest improve forecasts (Figure 5), but reduce the set of actionable items available before harvest.

There is a large diversity in methods, crops and study areas, but all with a bias towards staple crops and main producing countries. For many crop-country combinations there are well-performing methods. Yet, there is no silver bullet technique such that each new setting requires appropriate validation.

Major deficits remain as follows. Targeted research efforts might help improve upon them.

First, many regions and crops are under-represented (Figure S1). If forecasting yields is perceived as an adaptation mechanism for climate change, enhanced consideration of minor crops and areas is relevant to maintain a high diversity of nutrient sources. The gap is particularly prominent in developing countries, where yield forecasting is currently difficult due to small plot sizes, frequent mixed and inter-cropping, limited availability and resolution of weather or soil data and, finally, scarcity of crop yield data for validating methods.

Second, a rigorous out-of-sample validation (i.e. evaluation data is not used for model training) is lacking in more than half of the studies. This leaves the true performance of a method unknown, since operational forecasting is tantamount to out-of-sample prediction. There may also be a reporting bias for out-of-sample performance, since the latter is higher than the in-sample performance within our selected literature, when compared on average and not for individual studies – which is unexpected and may point to authors only reporting independent validation results if they are robust.

Third, model comparison and thus the choice of the 'best' model per region is only possible if they are compared under the same settings (input and output data). This is difficult due to mostly unavailable data from the studies.

Fourth, forecasting assumes that no major unexpected yield-changing events after the day of forecasting occur ('potential' yield outlook). In particular when no weather forecasts are included, a 'normal' weather is assumed for the rest of the growing season with no major biotic or abiotic stresses. Yet extreme events like floods or pest outbreaks can seriously change expectable harvest amounts, making their prediction relevant for yield forecasting (compare e.g. Ben-Ari et al. (2018) for the case in France 2016 or the locusts outbreaks in Eastern Africa and Pakistan in early 2020 (FAO, 2020)).

Fifth, there are several sources of uncertainty in yield forecasting which need to be reported appropriately but usually are not. The first type is due to the models applied, where parameters are uncertain and the choice of the best model is elusive. The second type of uncertainty is due to unobserved data during the current growing season, including actual cropping patterns, soil conditions, planting dates, management decisions like fertilization (Fieuzal and Baup, 2017; Stone and Meinke, 2005), and the spread of pests and diseases. The third type of uncertainty comes from the forecasted data that feed into the models (e.g. weather or pest distributions). Forecast distributions instead of point estimates are a step forward to clearly highlight ranges of outcomes. The uncertainty of forecasts may be decisive for policy planning.

Sixth, most of the studies on experimental plot level were applied to controlled conditions and may show reliable forecasting there. But their performance in real-world application, for larger areas, with less and possibly low-quality data, is usually not assessed.

Seventh, the reproducibility of studies is unclear. Correlations between input and yield data may be restricted to a certain period of time. For example, Cane et al. (1994) produced reliable maize yield forecasts in Zimbabwe from sea surface temperatures. Yet it is unclear whether this relation still holds, given changes in climate and cropping patterns, such that improvements or new approaches might be necessary.

Eighth, more complex crop models do not necessarily produce better results than simple models (Ben-Ari et al., 2016), but it is also apparent that statistical models do not fully reflect soil-plant-atmosphere interactions or the timing of stress (Mavromatis, 2016). Together, this suggests that a model ensemble approach with diverse model complexities might provide more reliable estimates. Few studies, though, have applied several methods to account for this, and a better capturing of extremes with model ensembles has yet to be proven.

Ninth, the spatial and temporal resolution of weather data and remote sensing products is mostly not high enough to represent diverse field conditions (Bolton and Friedl, 2013), in particular in developing countries with often cloudy conditions during large parts of the growing season. The trade-off between space and time is exemplified by the choice between MODIS (high revisit frequency) and Landsat or Sentinel satellites (higher spatial resolution). Similarly for weather data, small-scale local conditions may alter obtained precipitation and thus yield expectations, but are not reflected in most large-scale weather products. Crop surveys on the ground could amend estimates by providing local information, but are costly and time-intensive to obtain and thus difficult to upscale.

4.3. Strengths and weaknesses of weather data

Weather data, both historical and forecasted, are used in many crop forecasting methods, and have undergone strong developments for better availability, coverage and accuracy – but deficits related to crop forecasting remain. Despite continuous improvements of skill in numerical weather forecasts (Alley et al., 2019), the lack of skill after about 10 days (Bauer et al., 2015; Kushnir et al., 2019) is still a central bottleneck to any yield forecasting. There has been rapid progress at sub-seasonal and seasonal time scales (Klemm and McPherson, 2017). Yet these are still under-used, as their capabilities and limitations are not directly transparent to every user (Turco et al., 2017). For example, the Madden-Julian Oscillation (MJO) that wanders across the globe in 1-3 months, with important implications for the tropical summer monsoons and agriculture in general, can now be forecasted with predictive skills up to five weeks before onset (Kim et al., 2018). Experimental linkages to climate modes such as the ENSO index have the potential to improve the forecasting skill in certain years (Anderson et al., 2019). Implementing seasonal forecasting in agricultural outlooks can already positively affect farmers decision making (Gunda et al., 2017). Process-based crop models, in contrast to empirical models, usually require daily weather input, which are rarely available for long lead times or at high spatial resolution – the seasonal distribution of rainfall is the most critical information (Coelho and Costa, 2010). The influence of spatial resolution, though, has been described as limited when aggregating forecasts to larger scales (Wit et al., 2005). Capa-Morocho et al. (2016) evaluate methods to disaggregate seasonal forecasts anomalies into daily weather realizations for the CERES crop model in Spain with promising results. For all weather products a reduction in the lag time between observation and the provision of the data would be helpful to allow more timely forecasts. Finally, an improved forecasting of extreme events would aid to adjust yield expectations under non-linear influences.

4.4. Strengths and weaknesses of remote sensing data

A large range of satellites with different spatial and temporal resolutions is available to use for commercial or research purposes. Remote sensing data are increasingly used in yield forecasting, often with remarkable success in deducing harvest amounts from vegetation condition around anthesis. Since large parts of the growing season, for example in the tropics, can be in the rainy season, clouds may hinder the acquisition of useful images. Therefore, the combination of several resources is reasonable, although this necessitates harmonization across resolutions, wavelengths and indices. A recent effort in this direction is the Harmonized Landsat Sentinel archive (HLS; see results).

While the NDVI is most often applied, as a surrogate for green leaf area and plant health, other indices like the EVI, SAVI, NDWI or RGB channels also play a role (Cao et al., 2015), for example to limit soil and atmospheric disturbances (Kouadio et al., 2014) or to avoid NDVI saturation at large leaf areas (Peralta et al., 2016). Radar satellites have also been used for yield estimation, with considerable predictive skill for some crops (Fieuzal and Baup, 2017; Fieuzal et al., 2017). Radar signals are less disturbed by clouds than optical wavelengths such that a combination of radar and optical wavelengths might serve well during tropical rain-determined growing seasons. Moreover, most satellite-derived measures are useful until anthesis or the peak of green biomass accumulation, but diminishingly predictive around maturity. This precludes the harvest index (HI; the ratio of harvested to total biomass) to be estimated from space – but the HI is not constant over time or between cultivars and is an important determinant of yields (Fieuzal et al., 2017; Li et al., 2011; Walter et al., 2018). Therefore, the combination of several indices from different spectra and timepoints is advised to overcome limitations of each single index, exemplified by several studies in our database. The combination of vegetation measurements from satellites and unmanned aerial vehicles (UAV or drones) is another avenue of research that could overcome the resolution-revisit-coverage trade-off for selected areas.

Distinct methods for yield forecasting (process-based models, empirical models, expert integration or others) require – if they use remotely sensed data at all – different types of these data in terms of resolution, wavelengths, revisit frequency and accuracy, as has been reviewed by Basso and Liu (2019).

To increase the usefulness of satellites even further, a reduced lag time between data acquisition and provision would be helpful. In the tropics, though, the time between two cloud-free images may be more limiting than the time until data provision and calls for a merging of several satellite sources. Moreover, traits that are currently mostly measured manually (e.g. plant height, lodging, deep soil moisture, micronutrient content or pest infestation) could increasingly be measured from space, although the exact relationships between on-the-ground and remote measurements will have to be established first. A combination of remote sensing and weather data is advised (see also above) as all forecasts based on remote data alone cannot incorporate later yield-diminishing effects, which might be represented in weather forecasts. Finally, while remote sensing indicators deliver a comprehensive picture of vegetation status, it is often not possible to derive causes for a non-normal status (like pests, lack of water or nitrogen) and thus recommendations for management practices are difficult, if not augmented from other sources.

4.5. Strengths and weaknesses of available crop masks

To upscale locally robust forecasts of crop yields to larger scales, an accurate crop mask is recommended to avoid confounding from other crops grown in the same area. Crop masks are mostly required to derive area-dependent production estimates, but can also aid in improving yield forecasts in several ways. First, filtering irrelevant areas before forecasting can free up computing power for relevant areas and thus increase accuracy. Second, if yields are to be forecasted with the aid of remote sensing data, crop masks help to aggregate these only for the relevant pixels. Third, if empirical models are used to calculate the production (sometimes this information is can be derived earlier from import–export balances), exact crop masks can help to calculate the corresponding crop yields.

While the availability of weather or satellite data, as well as the capacity of computational methods, have increased recently, reliable crop masks that are available during the season are still not accessible. Thus current forecasts for larger regions either rely on crop masks from previous years or develop their own masking algorithm (19 studies in our database). Although several global or continental crop masks are available (Table S3), there are huge discrepancies between these in terms of resolution, coverage, update schedules and accuracy (on the latter aspect, cf. e.g. Lambert et al. (2016)). Most of the masks are also not crop specific, indicating only whether the pixel under scrutiny is predominantly crop or any other use. Crop masks are therefore considered a major uncertainty in large-scale forecasting of crop yields, in particular in evergreen areas or developing countries where cropping patterns are highly dispersed and subject to frequent changes (Vancutsem et al., 2012). If no reliable crop mask is available, yields (as harvest per area) can still be estimated based on weather data, in particular when growing areas are virtually constant over time.

4.6. Suggestions for improving methods for yield forecasting

Ideas for improving yield forecasting approaches are presented in this section. Other studies discussing amendment suggestions are, for example, Fritz et al. (2019), van der Velde et al. (2019) and Basso and Liu (2019), which we complement here.

4.6.1. Increase data availability and provide actionable forecasts

High-quality yield and management data are necessary to train and validate forecasting methods. While the USA is a paragon for open data, the situation is direr elsewhere, especially in developing nations. Data availability could be improved if forecasting studies usually publish their data, which could be done via separate data descriptor papers. Semi-public resources like social media or newspapers could be tapped by data mining approaches, which have already proven valuable in early warning systems (Ford et al., 2016), although their use is not straightforward (Palen and Anderson, 2016). Crowd-sourcing initiatives are increasingly getting in focus, for example for the Picture-Based Insurance (PBI) project (Ceballos et al., 2019) and could be of use for enhancing ground truth data with limited resources. The practice of crop cuts (where samples from pre-defined random fields are aggregated to represent average yields for a region; e.g. Murthy et al. (1997)), although sometimes criticized for its inaccuracy, time consumption and non-representativeness (Lobell et al., 2019), could be an option to increase data availability in all countries if properly performed. Data stemming from precision farming could be of potential use, for instance intra-field yield estimates from combine harvesters or, similarly, fertilizer or pesticide application. Finally, the choice of the growing season is decisive to capture key phenological phases and the timing of stressors. Yet in most regions only coarse season calendars are available, which hampers locally adapted forecasting efforts.

If forecasts should be useful for in-season adaptation or planning, they must provide actionable items. Different target groups have different requirements, illustrated with three examples. First, for farmers, the forecast should be available as early as possible on plot level to ensure that agronomic interventions (fertilization, irrigation, cultivar choice, etc.) are still possible. High accuracy is preferable, but farmers can also plan with tendency knowledge, that is whether the remaining part of the growing season will be dryer or wetter than usual. Guidelines can be taken, for example, from Frame et al. (2017) to translate forecasted losses into literate terms, using the triad of ‘unusual, unfamiliar, unknown’. Second, for commodity traders, potentially any lead time which allows financial adjustments is appreciated – yet the pre-harvest information is most valuable when not publicly available as this is a major benefit over other traders. A sufficiently accurate reflection of extreme production shocks is necessary, though, to allow for estimating non-linear impacts on commodity prices. Third, for humanitarian aid organizations, production forecasts are an indication, but these could be augmented by prospects on ensuing food availability, nutritional content or safety, as key ingredients of food security (Schmidhuber and Tubiello, 2007). Developing countries are main targets for aid, but often simultaneously scarce in data. Thus, only moderate accuracy may be expected, but then with high robustness and long lead time to mobilize funds and act. Further literature on the diverse needs of stakeholders indicates that it is only limitedly possible to unify forecasting needs (Bocca et al., 2015; Challinor, 2009; Guimarães Nobre et al., 2019; Hansen, 2002; Rijks and Baradas, 2000; van der Velde et al., 2019).

4.6.2. Improve yield forecasting methods on multiple levels

Regarding input data, improvement suggestions for weather and remote sensing data are illustrated above. Further factors like pests and diseases, management decisions, air pollution or economic circumstances (e.g. fertilizer prices) are equally relevant and could be attempted to be forecasted, using models or simple proxy forecasts. The combination of inputs from independent and diverse sources is recommended to increase robustness of forecasts and to better assess uncertainty.

Regarding models, ensemble approaches seem timely, as a multitude of models allows to select the best one for each region and thus to overcome idiosyncratic deficiencies. In any case, every model needs to be validated thoroughly, which is only possible with an out-of-sample assessment where evaluation data are invisible during the training process. A lens should be placed on extreme years (positive and negative) to capture yield drivers also in non-standard years, which may have highly non-linear repercussions for farmers and commodity markets (Headey, 2011). Long time frames for model training and validation are helpful here, as they contain more variation. A further enhancement in methods is to consider the phenology of crop growth and the timing of stress factors, as effects may differ during the season (Barnabas et al., 2008). Finally, adapting methods from similar (climate, soil, crop, management) regions across the globe may help to quickly assemble first forecasts for previously untapped regions. Our review aids in this respect by providing a comprehensive overview. Regarding results, a preference for robust, actionable forecasts over highly optimized quantitative forecasts (which may not hold in subsequent years) may suit selected purposes. Actionable also requires to clearly communicate uncertainty, for example how likely a large forecasted loss is across different input data. Potgieter et al. (2003) suggest probabilistic quality measures for crop yield forecasts.

4.6.3. Crop and region-dependent improvement options along with standardization suggestions

The strong bias in studied crops and regions (Figure 1, Figure 2, Figures S1-S2) urges to consider less researched cases. For many crop-country combinations, studies are present but may already be outdated (more than 10 years since publication; examples: maize in Ethiopia or rye in Germany) or have not yet been subject to a rigorous out-of-sample evaluation (examples: Bezuidenhout and Schulze (2006), Bognár et al. (2011)). There are also hundreds of cases where the FAO reports harvested areas of more than 10,000 ha but no forecast has yet been conducted in the scientific literature (examples: maize in India or potatoes in Hungary). Moreover, many studies are based on only few years of data – there are 76 studies with only one analysis year, particularly for experimental studies – and thus require longer time series to be reliably extrapolated to unknown conditions. It is also recommended to test the transfer of results from research plots onto larger geographic scales. Further examples on each of these deficits can be derived from Figures S1-2 and Table S2.

To achieve robust results in crop yield forecasting, we suggest the following steps for standardizing future studies, based on findings detailed above. These mainly relate to scientific studies, but can also serve in operational systems.

First, we suggest to consider multiple methods in forecasting. Statistical models, process-based models or machine-learning methods consider different aspects of crop growth and may thus complement each other. Second, forecast studies need to report out-of-sample results on independent test data (otherwise they are “not even wrong”, following Ghahramani (2015)). We also recommend to report method performance in at least one common evaluation measure, namely R^2 or RMSE to facilitate cross-study comparisons. Third, the acquisition of quality-checked input data is recommended. There are considerable sources of error in yields or management data which may impede good results. Gap-filling, filtering, substitution or imputation methods may cover up for data deficiencies. Fourth, if remote sensing data are used, we recommend the testing and inclusion of multiple sources and indices. In our literature database there are more than 60 examples how to select appropriate wavelengths and combine them. Fifth, the data used in studies should be made publicly available, either within the study or as a separate data descriptor paper to allow other researchers to reproduce results and test different methods on the same data set. Sixth, and finally, a systematic treatment of uncertainty within a Bayesian framework or model ensemble is recommended to deduce robustness of results. With these suggestions we hope to spur the development of more robust and reproducible forecasting methods that can then readily be included into operational systems.

5. Conclusion

We presented a systematic review of crop yield forecasting methods and three often used data domains: weather, remote sensing and crop masks. We conclude that yield forecasts are increasingly feasible for many crops and regions and that more input data resources have become available. Deficits on all areas remain, though, and should be targeted in research. Specifically, there are large crop growth areas with limited or no research on forecasting, uncertainties or unavailability of necessary data resources, method-specific deficiencies and lack of robust or coherent accuracy assessments across studies. To overcome these deficits, we suggest to target under-researched areas by transferring established approaches, increase data availability from published studies, combine several methods and data sources to unite their strengths, follow standardized procedures for designing forecasting studies and, finally, provide practically actionable forecasts in addition to scientific achievements. Reliable crop yield forecasting can critically improve planning of agronomic management and adaptation measures, stabilize farmers’ income and thus could become an integral component of food security early warning systems.

Author contributions

BS initiated, designed and performed the study and wrote the manuscript. JJ contributed to research design, literature analysis, wrote the review of forecast weather data and contributed to writing. CG contributed to research design and writing.

Acknowledgements

The authors thank Marijn van der Velde, Joshua Elliott and the participants of a workshop on near-term yield forecasting during the Impacts World 2017 conference in Potsdam for their valuable comments and ideas. BS acknowledges funding by Climate-KIC supported by the EIT. This work has been funded by the ClimSec Sahel (2518AA0291) project financed by the German Ministry for Foreign Affairs.

Conflicts of interest

None

References

Alley, R.B., Emanuel, K. and Zhang, F., 2019. Advances in weather prediction: Better weather and environmental forecasting will continue to improve well-being. Science, 363(6425): 342-4.

Anderson, W., You, L., Wood, S., Wood-Sichra, U. and Wu, W., 2015. An analysis of methodological and spatial differences in global cropping systems models and maps. Global Ecology and Biogeography, 24: 180-191.

Anderson, W.B., Seager, R., Baethgen, W., Cane, M. and You, L., 2019. Synchronous crop failures and climate-forced production variability. Science Advances, 5.

Anwar, M.R., Rodriguez, D., Liu, D.L., Power, S. and O'Leary, G.J., 2008. Quality and potential utility of ENSO-based forecasts of spring rainfall and wheat yield in south-eastern Australia. Australian Journal of Agricultural Research, 59(2).

Azzari, G. and Lobell, D.B., 2017. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. Remote Sensing of Environment, 202: 64-74.

Balaghi, R., Tychon, B., Eerens, H. and Jlibene, M., 2008. Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. International Journal of Applied Earth Observation and Geoinformation, 10(4): 438-452.

Barnabas, B., Jager, K. and Feher, A., 2008. The effect of drought and heat stress on reproductive processes in cereals. Plant, cell & environment, 31(1): 11-38.

Basso, B., Cammarano, D. and Carfagna, E., 2013. Review of Crop Yield Forecasting Methods and Early Warning Systems, First Meeting of the Scientific Advisory Committee of the Global Strategy to Improve Agricultural and Rural Statistics. FAO.

Basso, B. and Liu, L., 2019. Seasonal crop yield forecast: Methods, applications, and accuracies. Advances in Agronomy, 154: 201-255.

Bauer, P., Thorpe, A. and Brunet, G., 2015. The quiet revolution of numerical weather prediction. Nature, 525(7567): 47-55.

Ben-Ari, T. et al., 2016. Indentifying indicators of extreme wheat and maize yield losses. Agricultural and Forest Meteorology: 130-140.

Ben-Ari, T. et al., 2018. Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. Nature communications, 9(1): 1627.

Bernardi, M., Delincé, J., Durand, W., Zhang, N. and Fonteneau, F., 2016. Crop Yield Forecasting: Methodological and Institutional Aspects, FAO, Rome.

Bezuidenhout, C.N. and Schulze, R.E., 2006. Application of seasonal climate outlooks to forecast sugarcane production in South Africa. Climate Research, 30: 239-46.

Bilotta, G.S., Milner, A.M. and Boyd, I., 2014. On the use of systematic reviews to inform environmental policies. Environmental Science & Policy, 42: 67-77.

Bocca, F.F., Rodrigues, L.H.A. and Arraes, N.A.M., 2015. When do I want to know and why? Different demands on sugarcane yield predictions. Agricultural Systems, 135: 48-56.

Boers, N. et al., 2019. Complex networks reveal global pattern of extreme-rainfall teleconnections. Nature, 566(7744): 373-377.

Bognár, P. et al., 2011. Yield forecasting for wheat and corn in Hungary by satellite remote sensing. International Journal of Remote Sensing, 32(17): 4759-4767.

Bolton, D.K. and Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agricultural and Forest Meteorology, 173: 74-84.

Borovikov, A. et al., 2017. GEOS-5 seasonal forecast system. Climate Dynamics, 53(12): 7335-7361.

Brown, J.N., Hochman, Z., Holzworth, D. and Horan, H., 2018. Seasonal climate forecasts provide more definitive and accurate crop yield predictions. Agricultural and Forest Meteorology, 260-261: 247-254.

Cane, M., Eshel, G. and Bucklandt, R.W., 1994. Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea surface temperature. Nature, 370: 204-5.

Cao, Q. et al., 2015. Improving in-season estimation of rice yield potential and responsiveness to topdressing nitrogen application with Crop Circle active crop canopy sensor. Precision Agriculture, 17(2): 136-154.

Capa-Morocho, M. et al., 2016. Crop yield outlooks in the Iberian Peninsula: Connecting seasonal climate forecasts with crop simulation models. Agricultural Systems, 149: 75-87.

Ceballos, F., Kramer, B. and Robles, M., 2019. The feasibility of picture-based insurance (PBI): Smartphone pictures for affordable crop insurance. Development Engineering, 4.

Challinor, A., 2009. Towards the development of adaptation options using climate and crop yield forecasting at seasonal to multi-decadal timescales. Environmental Science & Policy, 12(4): 453-465.

Chipanshi, A. et al., 2015. Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agricultural and Forest Meteorology, 206: 137-150.

Chivasa, W., Mutanga, O. and Biradar, C., 2017. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: a review. International Journal of Remote Sensing, 38(23): 6816-6845.

Coelho, C.A. and Costa, S.M., 2010. Challenges for integrating seasonal climate forecasts in user applications. Current Opinion in Environmental Sustainability, 2(5-6): 317-25.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M. and Jones, P.D., 2018. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. Journal of Geophysical Research: Atmospheres, 123(17): 9391-9409.

Daly, C. et al., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology, 28(15): 2031-2064.

DayMet, 2020. https://daymet.ornl.gov/overview; accessed on Feb 18, 2020.

Delincé, J., 2017. Recent practices and advances for AMIS crop yield forecasting at farm and parcel level: A review, FAO, Rome.

ECMWF, 2020. https://www.ecmwf.int/en/forecasts/documentation-and-support/medium-range-forecasts; accessed on Feb 18, 2020.

Egelkraut, T.M., Garcia, P., Irwin, S.H. and Good, D.L., 2015. An Evaluation of Crop Forecast Accuracy for Corn and Soybeans: USDA and Private Information Agencies. Journal of Agricultural and Applied Economics, 35(1): 79-95.

El Jarroudi, M. et al., 2012. Integrating the impact of wheat fungal diseases in the Belgian crop yield forecasting system (B-CYFS). European Journal of Agronomy, 40: 8-17.

Elavarasan, D., Vincent, D.R., Sharma, V., Zomaya, A.Y. and Srinivasan, K., 2018. Forecasting yield by integrating agrarian factors and machine learning models: A survey. Computers and Electronics in Agriculture, 155: 257-282.

ERA5, 2020. https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation; accessed on Feb 18, 2020.

FAO, 2017. Building Agricultural Market Information Systems: A literature review, FAO, Rome.

FAO, 2020. http://www.fao.org/ag/locusts/en/info/info/index.html; accessed on 11 March 2020.

Ferrise, R. et al., 2015. Monthly-to-seasonal predictions of durum wheat yield over the Mediterranean Basin. Climate Research, 65: 7-21.

FEWS.NET, 2020. https://fews.net/; accessed on Feb 03, 2020.

Fieuzal, R. and Baup, F., 2017. Forecast of wheat yield throughout the agricultural season using optical and radar satellite images. International Journal of Applied Earth Observation and Geoinformation, 59: 147-56.

Fieuzal, R., Marais Sicre, C. and Baup, F., 2017. Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. International Journal of Applied Earth Observation and Geoinformation, 57: 14-23.

Ford, J.D. et al., 2016. Opinion: Big data has big potential for applications to climate change adaptation. Proceedings of the National Academy of Sciences of the United States of America, 113(39): 10729-32.

Frame, D., Joshi, M., Hawkins, E., Harrington, L.J. and de Roiste, M., 2017. Population-based emergence of unfamiliar climates. Nature Climate Change, 7(6): 407-411.

Fritz, S. et al., 2019. A comparison of global agricultural monitoring systems and current gaps. Agricultural Systems, 168: 258-272.

Fritz, S., See, L. and Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. International Journal of Remote Sensing, 31(9): 2237-2256.

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

Funk, C. et al., 2019a. A High-Resolution 1983–2016 Tmax Climate Data Record Based on Infrared Temperatures and Stations by the Climate Hazard Center. Journal of Climate, 32(17): 5639-5658.

Funk, C. et al., 2019b. Recognizing the Famine Early Warning Systems Network: Over 30 Years of Drought Early Warning Science Advances and Partnerships Promoting Global Food Security. Bulletin of the American Meteorological Society, 100(6): 1011-1027.

Gelaro, R. et al., 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J Clim, Volume 30(Iss 13): 5419-5454.

GEOGLAM, 2020. http://earthobservations.org/geoglam.php; also https://cropmonitor.org/; both accessed on Feb 03, 2020.

GEWEX, 2020. https://gewex-srb.larc.nasa.gov/common/php/SRB_about.php and https://ceres.larc.nasa.gov/products.php?product=FLASHFlux-Level2; accessed on Feb 18, 2020.

GFS, 2020. https://www.emc.ncep.noaa.gov/index.php?branch=GFS; accessed on Feb 18, 2020.

Ghahramani, Z., 2015. Probabilistic machine learning and artificial intelligence. Nature, 521(7553): 452-459.

Global Commission on Adaptation, 2019. https://cdn.gca.org/assets/2019-09/GlobalCommission_Report_FINAL.pdf; accessed on 11 March 2020.

Grekousis, G., Mountrakis, G. and Kavouras, M., 2015. An overview of 21 global and 43 regional land-cover mapping products. International Journal of Remote Sensing, 36(21): 5309-5335.

Guimarães Nobre, G. et al., 2019. Financing agricultural drought risk through ex-ante cash transfers. Science of The Total Environment, 653: 523-535.

Gunda, T., Bazuin, J.T., Nay, J. and Yeung, K.L., 2017. Impact of seasonal forecast use on agricultural income in a system with varying crop costs and returns: an empirically-grounded simulation. Environmental Research Letters, 12(3).

Guo, W.W. and Xue, H., 2011. An incorporative statistic and neural approach for crop yield modelling and forecasting. Neural Computing and Applications, 21(1): 109-117.

Hansen, J.W., 2002. Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. Agricultural Systems, 74(3): 309-330.

Headey, D., 2011. Rethinking the global food crisis: The role of trade shocks. Food Policy, 36(2): 136-146.

HLS, 2020. https://hls.gsfc.nasa.gov/; accessed on Feb 07, 2020.

Huffman, G.J. et al., 2019. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG).

Iizumi, T., Shin, Y., Kim, W., Kim, M. and Choi, J., 2018. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. Climate Services, 11: 13-23.

Inglada, J. et al., 2015. Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. Remote Sensing, 7(9): 12356-12379.

Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sensing of Environment, 141: 116-128.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996. The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorlogical Society, 77: 437-472.

Kamali, B., Abbaspour, K.C., Lehmann, A., Wehrli, B. and Yang, H., 2018. Spatial assessment of maize physical drought vulnerability in sub-Saharan Africa: Linking drought exposure with crop failure. Environmental Research Letters, 13(7).

Kastens, J. et al., 2005. Image masking for crop yield forecasting using AVHRR NDVI time series imagery. Remote Sensing of Environment, 99(3): 341-356.

Kim, H., Vitart, F. and Waliser, D.E., 2018. Prediction of the Madden–Julian Oscillation: A Review. Journal of Climate, 31(23): 9425-9443.

Klemm, T. and McPherson, R.A., 2017. The development of seasonal climate forecasting for agricultural producers. Agricultural and Forest Meteorology, 232: 384-399.

Knippertz, P., Christoph, M. and Speth, P., 2003. Long-term precipitation variability in Morocco and the link to the large-scale circulation in recent and future climates. Meteorology and Atmospheric Physics, 83(1-2): 67-88.

Koirala, A., Walsh, K.B., Wang, Z. and McCarthy, C., 2019. Deep learning – Method overview and review of use for fruit detection and yield estimation. Computers and Electronics in Agriculture, 162: 219-234.

Kouadio, L., Newlands, N., Davidson, A., Zhang, Y. and Chipanshi, A., 2014. Assessing the Performance of MODIS NDVI and EVI for Seasonal Crop Yield Forecasting at the Ecodistrict Scale. Remote Sensing, 6(10): 10193-10214.

Kushnir, Y. et al., 2019. Towards operational predictions of the near-term climate. Nature Climate Change, 9(2): 94-101.

Lambert, M.-J., Waldner, F. and Defourny, P., 2016. Cropland Mapping over Sahelian and Sudanian Agrosystems: A Knowledge-Based Approach Using PROBA-V Time Series at 100-m. Remote Sensing, 8(3): 232.

Lehmann, J., Kretschmer, M., Schauberger, B. and Wechsung, F., 2020. Potential for early forecast of Moroccan wheat yields based on climatic drivers. Geophysical Research Letters, n/a(n/a): e2020GL087516.

Li, H. et al., 2011. A comparison of harvest index estimation methods of winter wheat based on field measurements of biophysical and spectral data. Biosystems Engineering, 104(4): 396-403.

Liu, L. and Basso, B., 2020. Linking field survey with crop modeling to forecast maize yield in smallholder farmers' fields in Tanzania. Food Security.

Lobell, D.B. et al., 2019. Sight for Sorghums: Comparisons of Satellite- and Ground-Based Sorghum Yield Estimates in Mali. Remote Sensing, 12(1).

López-Lozano, R. et al., 2015. Towards regional grain yield forecasting with 1km-resolution EO biophysical products: Strengths and limitations at pan-European level. Agricultural and Forest Meteorology, 206: 12-32.

MacDonald, R.B. and Hall, F.G., 1980. Global Crop Forecasting. Science, 208: 670-679.

Mahalanobis Centre, 2020. NCFC: https://www.ncfc.gov.in/ (acc. Feb 03, 2020).

Mahon, N., Crute, I., Simmons, E. and Islam, M.M., 2017. Sustainable intensification – "oxymoron" or "third-way"? A systematic review. Ecological Indicators, 74: 73-97.

Makowski, D., Nesme, T., Papy, F. and Doré, T., 2013. Global agronomy, a new field of research. A review. Agronomy for Sustainable Development, 34(2): 293-307.

Maresma, Á., Ariza, M., Martínez, E., Lloveras, J. and Martínez-Casasnovas, J., 2016. Analysis of Vegetation Indices to Determine Nitrogen Application and Yield Prediction in Maize (Zea mays L.) from a Standard UAV Service. Remote Sensing, 8(12).

Mavromatis, T., 2016. Spatial resolution effects on crop yield forecasts: An application to rainfed wheat yield in north Greece with CERES-Wheat. Agricultural Systems, 143: 38-48.

McKenzie, A.M., 2008. Pre-Harvest Price Expectations for Corn: The Information Content of USDA Reports and New Crop Futures. American Journal of Agricultural Economics, 90(2): 351-366.

Murthy, C.S., Thiruvengadachari, S., Jonna, S. and Raju, P.V., 1997. Design of crop cutting experiments with satellite data for crop yield estimation in irrigated command areas. Geocarto International, 12(2): 5-11.

NASA POWER, 2020. https://power.larc.nasa.gov/, accessed on Feb 18, 2020.

Newlands, N.K. et al., 2014. An integrated, probabilistic model for improved seasonal forecasting of agricultural crop yield under environmental uncertainty. Frontiers in Environmental Science, 2.

NLDAS, 2020. https://ldas.gsfc.nasa.gov/nldas; accessed on Feb 18, 2020.

NOAA CPC, 2020. https://www.cpc.ncep.noaa.gov/; accessed on Feb 18, 2020.

Palen, L. and Anderson, K.M., 2016. Crisis informatics—New data for extraordinary times. Science, 353(6296): 224-5.

Peralta, N., Assefa, Y., Du, J., Barden, C. and Ciampitti, I., 2016. Mid-Season High-Resolution Satellite Imagery for Forecasting Site-Specific Corn Yield. Remote Sensing, 8(10): 848.

Pervez, M.S. and Brown, J.F., 2010. Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics. Remote Sensing, 2(10): 2388-2412.

Potgieter, A.B., Everingham, Y.L. and Hammer, G.L., 2003. On measuring quality of a probabilistic commodity forecast for a system that incorporates seasonal climate forecasts. International Journal of Climatology, 23(10): 1195-1210.

Puma, M.J., Bose, S., Chon, S.Y. and Cook, B.I., 2015. Assessing the evolving fragility of the global food system. Environmental Research Letters, 10(2): 024007.

Rahmstorf, S. and Coumou, D., 2011. Increase of extreme events in a warming world. Proceedings of the National Academy of Sciences of the United States of America, 108(44): 17905-9.

Rembold, F., Atzberger, C., Savin, I. and Rojas, O., 2013. Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection. Remote Sensing, 5(4): 1704-1733.

Rijks, D. and Baradas, M.W., 2000. The clients for agrometeorological information. Agricultural and Forest Meteorology, 103(1): 27-42.

Rocha, H. and Dias, J.M., 2019. Early prediction of durum wheat yield in Spain using radial basis functions interpolation models based on agroclimatic data. Computers and Electronics in Agriculture, 157: 427-435.

Rodell, M. et al., 2004. The global land data assimilation system. Bulletin of the American Meteorlogical Society: 381-394.

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. Agricultural and Forest Meteorology, 200: 233-248.

S2S, 2020. http://www.s2sprediction.net/; accessed on Feb 18, 2020.

Saha, S. et al., 2014. The NCEP Climate Forecast System Version 2. Journal of Climate, 27(6): 2185-2208.

Saha, S., S. Moorthi, H. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y. Hou, H. Chuang, H.H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. van den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R.W. Reynolds, G. Rutledge, and M. Goldberg, 2010. The NCEP Climate Forecast System Reanalysis. Bulletin of the American Meteorlogical Society, 91: 1015-1057.

Samaniego, L. et al., 2018. Anthropogenic warming exacerbates European soil moisture droughts. Nature Climate Change, 8(5): 421-426.

Schmidhuber, J. and Tubiello, F.N., 2007. Global food security under climate change. Proceedings

of the National Academy of Sciences of the United States of America, 104(50): 19703-8.
Sheffield, J. and Wood, E.F., 2008. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. Climate Dynamics, 31(1): 79-105.
Stone, R.C. and Meinke, H., 2005. Operational seasonal forecasting of crop performance. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 360(1463): 2109-24.
Stott, P., 2016. How climate change affects extreme weather events. Science, 352(6293): 1517-1518.
Stratoulias, D. et al., 2017. A Workflow for Automated Satellite Image Processing: from Raw VHSR Data to Object-Based Spectral Information for Smallholder Agriculture. Remote Sensing, 9(10): 1048.
Takaya, Y. et al., 2016. Japan Meteorological Agency/Meteorological Research Institute-Coupled Prediction System version 1 (JMA/MRI-CPS1) for operational seasonal forecasting. Climate Dynamics, 48(1-2): 313-333.
TIGGE, 2020. https://confluence.ecmwf.int/display/TIGGE; accessed on Feb 18, 2020.
Tompkins, A.M. and al, e., 2017. The Climate-System Historical Forecast Project: Providing Open Access to Seasonal Forecast Ensembles from Centers around the Globe. Bulletin of the American Meteorlogical Society, 98: 2293-2301.
Turco, M. et al., 2017. Summer drought predictability over Europe: empirical versus dynamical forecasts. Environmental Research Letters, 12(8).
Tychon, B. et al., 2003. Belgian and Chinese Crop Growth Monitoring Systems Comparison, Adaptation and Improvement - An introduction. In: B. Tychon (Editor), Belgian and Chinese Crop Growth Monitoring Systems Comparison, Adaptation and Improvement. Fondation Universitaire Luxembourgeoise, Arlon, Belgium, pp. 135.
Valero, S. et al., 2016. Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions. Remote Sensing, 8(1): 55.
van der Velde, M. et al., 2018. In-season performance of European Union wheat forecasts during extreme impacts. Sci Rep, 8(1): 15420.
van der Velde, M. et al., 2019. Use and relevance of European Union crop monitoring and yield forecasts. Agricultural Systems, 168: 224-230.
Vancutsem, C., Marinho, E., Kayitakire, F., See, L. and Fritz, S., 2012. Harmonizing and Combining Existing Land Cover/Land Use Datasets for Cropland Area Monitoring at the African Continental Scale. Remote Sensing, 5(1): 19-41.
Vitart, F. and Robertson, A.W., 2018. The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. npj Climate and Atmospheric Science, 1(1).
Waldner, F. et al., 2017. National-scale cropland mapping based on spectral-temporal features and outdated land cover information. PLoS One, 12(8): e0181911.
Walter, J., Edwards, J., McDonald, G. and Kuchel, H., 2018. Photogrammetry for the estimation of wheat biomass and harvest index. Field Crops Research, 216: 165-174.
White, J.W., Hoogenboom, G., Kimball, B.A. and Wall, G.W., 2011. Methodologies for simulating impacts of climate change on crop production. Field Crops Research, 124(3): 357-368.
Wit, A.J.W.d., Boogaard, H.L. and Diepen, C.A.v., 2005. Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. Agricultural and Forest Meteorology, 135(1-4): 156-168.
Wu, B. et al., 2013. Remote sensing-based global crop monitoring experiences with China's CropWatch system. International Journal of Digital Earth, 7(2): 1-25.
Yang, Q., Shi, L., Han, J., Zha, Y. and Zhu, P., 2019. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. Field Crops Research, 235: 142-153.
Zhong, L., Hu, L. and Zhou, H., 2019. Deep learning based multi-temporal crop classification. Remote Sensing of Environment, 221: 430-443.

1122