Agricultural and Forest Meteorology 200 (2015) 233–248

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation

Alex C. Ruane[a,*], Richard Goldberg[b], James Chryssanthacopoulos[b]

[a] NASA Goddard Institute for Space Studies, New York, NY, United States
[b] Columbia University Center for Climate Systems Research, New York, NY, United States

ARTICLE INFO

Article history:
Received 31 March 2014
Received in revised form
18 September 2014
Accepted 22 September 2014

Keywords:
Climate forcing data
Daily climate series
Climate change
AgMIP
Climate impacts
Agriculture

ABSTRACT

The AgMERRA and AgCFSR climate forcing datasets provide daily, high-resolution, continuous, meteorological series over the 1980–2010 period designed for applications examining the agricultural impacts of climate variability and climate change. These datasets combine daily resolution data from retrospective analyses (the Modern-Era Retrospective Analysis for Research and Applications, MERRA, and the Climate Forecast System Reanalysis, CFSR) with in situ and remotely-sensed observational datasets for temperature, precipitation, and solar radiation, leading to substantial reductions in bias in comparison to a network of 2324 agricultural-region stations from the Hadley Integrated Surface Dataset (HadISD). Results compare favorably against the original reanalyses as well as the leading climate forcing datasets (Princeton, WFD, WFD-EI, and GRASP), and AgMERRA distinguishes itself with substantially improved representation of daily precipitation distributions and extreme events owing to its use of the MERRA-Land dataset. These datasets also peg relative humidity to the maximum temperature time of day, allowing for more accurate representation of the diurnal cycle of near-surface moisture in agricultural models. AgMERRA and AgCFSR enable a number of ongoing investigations in the Agricultural Model Intercomparison and Improvement Project (AgMIP) and related research networks, and may be used to fill gaps in historical observations as well as a basis for the generation of future climate scenarios.

Published by Elsevier B.V.

1. Introduction

The Agricultural Model Intercomparison and Improvement Project (AgMIP; Rosenzweig et al., 2013a) is conducting a wide range of climate-impacts-oriented activities focusing on crop and livestock models at the local level (e.g., Asseng et al., 2013; Singels et al., 2013; Bassu et al., 2014; Li et al., 2014; Ruane et al., 2014b) and on a global grid (Rosenzweig et al., 2013b), regional assessments of food security (Rosenzweig et al., 2012), and global economic impacts (e.g., Nelson et al., 2013; von Lampe et al., 2014). Related regional research networks such as the Consultative Group on International Agricultural Research (CGIAR) Climate Change, Agriculture and Food Security (CCAFS) and MACSUR (Modeling European Agriculture with Climate Change for Food Security; Rötter et al., 2013) are dealing with similar tasks. Consistency and transparency in climate data and methods facilitate comparisons across regions or between models in each of these assessments, particularly when market linkages between regions are emphasized. In particular, recent advances in porting agricultural models for parallel processing on high-performance computing has dramatically increased the demand for global climate datasets capable of driving global gridded crop models (Rosenzweig et al., 2013b). The historical period is of primary and urgent interest, as data from recent years may be used to calibrate models and serve as the basis for the development of future climate scenarios using different statistical methods (Wilby et al., 2004).

Here we describe the development of two new climate forcing datasets (AgMERRA and AgCFSR) designed to meet the needs of AgMIP and similar agricultural impacts assessments (White et al., 2011a). As opposed to strictly climatic datasets, particular consideration is given to agricultural areas and the climatic factors that crops are known to respond to, including biases in mean growing season temperature and precipitation, the seasonal cycle, interannual variability, the frequency and sequence of rainfall events, and the distribution of sub-seasonal extremes.

The root of all climate forcing datasets is the network of in situ meteorological observations maintained by meteorological agencies around the world. The density and quality of these stations varies widely through space and time, with the best coverage in developed countries and less reliable coverage in the Tropics and

* Corresponding author. Tel.: +1 212 678 5640; fax: +1 212 678 5648.
E-mail address: alexander.c.ruane@nasa.gov (A.C. Ruane).

http://dx.doi.org/10.1016/j.agrformet.2014.09.016
0168-1923/Published by Elsevier B.V.

Table 1
Overview of Climate Forcing Datasets, including the AgMERRA and AgCFSR datasets introduced here. Highest resolution is the resolution at which the data are archived and most finely distinguishable, although for some variables multiple grid boxes may be given the same value as the effective resolution is more coarse.

Climate Forcing Dataset	Reference	Time period	Highest resolution	Reanalysis basis (and resolution)	Monthly target for temperature and precipitation
Princeton	Sheffield et al. (2006)	1948–2008	0.5° × 0.5°	Reanalysis-1 (∼2°)	CRU TS2.0, with corrections for high-latitude precipitation using GPCP and TRMM
WFD		1958–2001	0.5° × 0.5°	ERA-40 (1°)	CRU TS2.1 and GPCCv4 versions
WFD-EI	Weedon et al. (2012)	1979–2009	0.5° × 0.5°	ERA-Interim (0.4°)	CRU TS3.1 and GPCCv5/6 versions
GRASP	Iizumi et al. (2014)	1961–2010	1.125° × 1.125°	JRA25 (1.125°) and ERA-40 (2.5° version)	CRU TS3.10.01, time-constant correction factors derived from 1961 to 1990.
AgMERRA	This study	1980–2010	0.25° × 0.25°	MERRA (0.5° × 0.67°)	Blend of in situ (CRU TS3.1, GPCCv6, WM) and satellite (TRMM, CMORPH, PERSIANN) products
AgCFSR	This study	1980–2010	0.25° × 0.25°	CFSR (∼0.3°)	Blend of in situ (CRU TS3.1, GPCCv6, WM) and satellite (TRMM, CMORPH, PERSIANN) products

Southern Hemisphere (Lorenz and Kunstmann, 2012). These data are also not always accessible and transparent as they may require high acquisition fees, restrictive limitations on use, or additional processing and quality control beyond the scope of many agricultural modelers. Several groups have collected these data and constructed harmonized, global gridded datasets at monthly resolution (New et al., 2002; Schneider et al., 2011; Willmott and Matsuura, 1995; Hijmans et al., 2005), however these require weather generators to synthesize daily resolution before they may be applied to crop models and are therefore likely to miss events that are important to the calibration and validation of agricultural models. Regional gridded observational networks have also been created (e.g., E-Obs in Europe, Haylock et al., 2008; APHRODITE in Asia, Yatagai et al., 2012; CPC US Unified Precipitation, Higgins et al., 2000), however many regions and variables are not covered by any such network and intercomparing sites between regions with different methodologies introduces inconsistencies.

The overall meteorological observational network is larger than just stations, as weather balloons and airborne instruments provide information about the upper atmosphere and satellite-based observations (particularly beginning in the late 1970s and including direct estimates of precipitation since the late 1990s) augment the entire network. The atmospheric modeling community has developed retrospective-analyses (reanalyses) that assimilate all available state observations into a physically-consistent atmospheric model that utilizes atmospheric structure and dynamics to estimate spatial and variable gaps in the observations. These reanalyses were designed for process studies, emphasizing atmospheric structure and circulation over some impacts-relevant variables. Flux variables, such as precipitation and radiation, are modeled rather than assimilated. Additionally, 2-m temperature, wind speed, and humidity measurements are not assimilated, as reanalyses rely instead on balloon (rawinsonde) networks to assimilate in the free atmosphere and then model boundary–layer profiles. The adherence to physical principles can lead to biases even at assimilated locations where limitations in model parameterizations or spatial resolution cannot be overcome.

In an effort to correct some of the most glaring shortcomings of the reanalyses, the land-surface hydrology community led the development of climate forcing datasets that adjust the reanalyses' daily time series to match the monthly gridded climate datasets. This can prevent full closure of the water and energy cycles, but maintains many of the most important properties for impacts assessment. Schwalm et al. (2014) found that hydrologic models are quite sensitive to the selection of a climate forcing dataset in the US, but only recently has the same question been asked of the agricultural models (e.g., Ruane et al., 2014a; Iizumi et al., 2014) despite the fact that agricultural models do not have the benefit of aggregating potentially compensating errors across watersheds. Adam et al. (2006) note that many global gridded climate datasets are biased toward the populated areas where stations have been set up rather than the mountains surrounding these, for example. This bias may be problematic for hydrologic catchments, but likely benefits agricultural applications as farmlands tend to be in the valleys and plains that are overrepresented.

This paper presents two new climate forcing datasets developed for agricultural applications utilizing a newer generation of reanalyses that are not currently associated with any climate forcing dataset. These reanalyses' higher spatial resolution, improved model physics, and additional sources of assimilated data hold great potential for improved agroclimatic assessment. Section 2 describes the datasets used in the construction, calibration, and evaluation of the AgMERRA and AgCFSR climate forcing datasets. Section 3 details the specifications of these new datasets and provides the complete methodology for their generation. Section 4 compares AgMERRA and AgCFSR against observations, the original reanalyses that they are drawn from, and existing climate forcing datasets. Following a discussion of the datasets' strengths and weaknesses, we describe the potential for gap-filling applications. Finally, we provide conclusions and next steps in the development, extension, and application of climate forcing datasets for agricultural modeling.

2. Datasets

2.1. Climate datasets

2.1.1. Existing climate forcing datasets

Methodologies for the development of the AgMIP climate forcing datasets was motivated by similar climate forcing datasets developed for various applications in recent years (Table 1), with the hopes that that new datasets could provide dramatically improved sub-monthly weather characteristics and radiation data that would improve agricultural modeling. The Princeton Climate Forcing Dataset (Sheffield et al., 2006) was developed for hydrologic applications, deriving its daily time series from the National Centers for Environmental Prediction/National Center for Atmospheric Research Reanalysis-1 (Kalnay et al., 1996) and adjusting to match