SoftwareX 7 (2018) 15–22

Contents lists available at ScienceDirect

SoftwareX

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

Original software publication

Web-based access, aggregation, and visualization of future climate projections with emphasis on agricultural assessments

Nelson B. Villoria [a,*], Joshua Elliott [b], Christoph Müller [c], Jaewoo Shin [d], Lan Zhao [e], Carol Song [e]

[a] Department of Agricultural Economics, Kansas State University, 331i Waters Hall, Manhattan, KS, 66506, USA
[b] University of Chicago and ANL Computation Institute Center for Robust Decision-making in Climate and Energy Policy (RDCEP) 5735 S. Ellis Ave., Chicago, IL 60637, USA
[c] Potsdam Institute for Climate Impact Research, Telegraphenberg A 31 14473, Potsdam, Germany
[d] Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA
[e] Rosen Center for Advanced Computing, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 26 September 2016
Received in revised form 27 October 2017
Accepted 14 November 2017

Keywords:
Climate data
CMIP5
Climate models
HUBzero
Growing season temperature
Growing season precipitation
Climate change and agriculture

ABSTRACT

Access to climate and spatial datasets by non-specialists is restricted by technical barriers involving hardware, software and data formats. We discuss an open-source online tool that facilitates downloading the climate data from the global circulation models used by the Inter-Sectoral Impacts Model Intercomparison Project. The tool also offers temporal and spatial aggregation capabilities for incorporating future climate scenarios in applications where spatial aggregation is important. We hope that streamlined access to these data facilitates analysis of climate related issues while considering the uncertainties derived from future climate projections and temporal aggregation choices.

Code metadata

Current code version	Version 1.0.1
Permanent link to code/repository used for this code version	https://github.com/ElsevierSoftwareX/SOFTX_SOFTX-D-16-00082
Legal Code License	GNU General Public License
Code versioning system used	LZ
Software code languages, tools, and services used	Java, R and Python, iData, Globus Online
Compilation requirements, operating environments & dependencies	Best on Debian GNU/Linux 7 with JDK 1.6, R 3.1.1 and Python 2.7.3 & Hub submit tool (isimiptransfer) needs to be installed
If available Link to developer documentation/manual	https://mygeohub.org/resources/1247/download/ClimateScenarioAggregator_UserManual.pdf
Support email for questions	Built-in HUBzero ticket support system

1. Motivation and significance

Studies of the effects of climate change on agriculture typically involve using observational data to determine the parameters connecting climate variables to agricultural productivity and then using future climate projections from global circulation models (GCM) to evaluate potential future impacts or the effects of alternative policies (e.g., [1]). Given the uncertainty surrounding future climate projections, it is considered best practice to use the output of several GCM in order to obtain a range of potential outcomes [2]. Coordination among climate modeling groups through the Coupled Model Intercomparison Project Phase 5 (CMIP5) and their collaboration with the Intergovernmental Panel on Climate Change, have greatly increased the availability of climate data. Yet, access by non-specialists is hindered by technical barriers. These

* Corresponding author.
E-mail address: nvillori@ksu.edu (N.B. Villoria).

https://doi.org/10.1016/j.softx.2017.11.004
2352-7110/

barriers include software, hardware, and the need for specialized skills to handle non-standard formats [3]. In addition to access to the data, the spatial processing is not trivial as it requires expertise in geographic information systems (GIS) methods to process both the climate data and the auxiliary datasets [4].

2. Scientific rationale

The tool discussed in this article is designed to reduce the technical barriers to accessing climate model outputs. For this, we have built a web-based tool that facilitates downloading and aggregating global grids (0.5 degree) of bias-corrected, historical and future monthly mean temperature and precipitation from the five General Circulation Models (GCMs) used by the Inter-Sectoral Impacts Model Intercomparison Project (ISI-MIP) [5,6]. (See Table 1 for the included models). The scientific problem the tool contributes to solving is how to facilitate the analysis of future climate scenarios in applications where spatial aggregation is important. This includes a wide range of economic analyses focused on either impact assessment [7–9] or policy analysis [10].

Access to the data produced by GCMs is far from trivial. For instance, bulk downloads from the Earth System Grid Federation,[1] an open data repository for the CMIP5, requires using Linux in order to run the bash script provided by the system. In addition, the bandwidth and storage capacity for the climate data are often limiting factors, both because of the size of the climate datasets as well as the number of existing models and emission scenarios. Moreover, manipulation of the spatio-temporal grids of the climate data requires considerable dexterity in using GIS software. This is even more so when the data needs to be merged with other datasets for aggregation over time and/or space. Procedures performed in user friendly, point and click interfaces of available GIS software are difficult to document, reproduce, and share. Best practices for handling spatial data involve using processing scripts, which requires knowledge of some more general-purpose programming languages.

In our experience, none of these barriers is necessarily a significant issue for users who have advanced training in climatology or computer science. However, for cross-disciplinary work, a lack of skills for accessing and processing climate data can be a major obstacle. In a widely read blog specializing in interactions between the environment and society, Auffhammer [16] addresses the difficulties economists face in trying to access the climate data from the CMIP5 archive. Hertel et al. [17] identify barriers to access to geo-referenced data as a main factor impeding a better understanding of how global environmental changes affect the sustainability of the global food system. To foster multidisciplinary work, geoprocessing on-line tools have been identified as being effective [18]. Such tools present a number of advantages including reducing software and hardware costs [19], leveraging shared cyberinfrastructure via web services [20], and combining elements of various workflows across different studies [21].

3. Target users

The tool targets mainly, but not exclusively, researchers interested in the effects of climate change on agriculture, but who lack the training and/or resources to obtain climate data projections. The target users of the Climate Scenario Aggregator (CSA) were identified in the course of a multi-year pilot effort originating in a request from the UK Foresight Programme to review the adequacy of the global data base infrastructure for analyzing issues related to agriculture and the environment [17]. The need for online tools to deliver large and complex geo-referenced datasets arose from an in-depth diagnosis of the availability of geospatial data for analyzing the impacts of global environmental change [21]. Moreover, the need for these tools was validated through three international workshops[2] with researchers and policy analysts working on climate change issues in both developed and developing countries. We also drew from our experience training graduate students in agricultural economics and computer science to work on multidisciplinary teams on issues related to climate change and global food security.

At the most general level, the CSA tool can be used as a downloading platform for the original GCM data in the ISI-MIP archive. The target user of this functionality is expected to be skilled in NetCDF formats, have a relatively powerful computer with reasonable bandwidth, and be comfortable with the scripting and/or programming languages needed for manipulating and processing spatially-explicit data. A second target user may need some assistance with basic preprocessing of the data, such as temporal and spatial aggregation. This user will benefit from the aggregation programs as well as the preprocessed datasets for temporal aggregation (crop calendars) and spatial aggregation (e.g., from grid cells to countries.) Finally, a third target user may be interested in the download and aggregation capabilities of the tool while employing alternative spatial aggregation schemes (e.g., gridded population.)

The CSA tool is related to other tools that seek to simplify access to (and spatial geoprocessing of) climate data while leveraging shared resources and expertise. For example, Wang et al. [3] developed user-friendly software applications for downscaling climate data for ecological modeling applications. Meanwhile, Villoria et al. [22] built an aggregation tool that facilitates access to the gridded projections of yield changes produced by the Agricultural Model Intercomparison and Improvement Project (AgMIP) [23]. This tool is being widely used as documented in usage logs and[3] peer-reviewed published articles [e.g., 10,24]. The CSA documented in this article has had a rapid uptake by the research community[4] and its range of applications is much more general than that of the AgMIP tool, so we expect it to have a larger impact.

4. Software description

The CSA tool is available at the GEOSHARE web site[5] and can be accessed using any standard Internet browser. This tool allows users to calculate, for each half-degree land pixel, a crop-specific growing season average value of temperature and precipitation using the global crop calendars from [25] (See Table 1 for crop coverage.) The tool also permits aggregating the pixels to large geographic units using crop harvested area and production from [26]. All of the source code – a Java graphical user interface (GUI) and a set of R functions – can be freely downloaded from the tool's landing page. The documentation and support for users include a User's Manual as well as a set of default regional maps and weighting schemes.

1 https://cds.nccs.nasa.gov/tools-services/esgf/.

2 Two workshops were held at Purdue University, West Lafayette, Indiana, on May 23, 2011 (https://mygeohub.org/groups/geoshare/workshop2011) and on September 10, 2014 [21]. The third workshop was held at the International Livestock Research Institute in Addis Ababa, Ethiopia, on March 11–15 2013, http://www.cgiar-csi.org/meetings/africa-agriculture-gis-week-2013.

3 The tool is available at https://mygeohub.org/tools/agmip. As of July 06 2017, the tool had 139 registered users who had performed more than 20 thousand data downloads.

4 The tool has been online since August 2015. As of July 06 2017, there are 30 registered users that have performed 343 data downloads. These statistics exclude the development team. An update usage log is available at https://mygeohub.org/resources/climatetool/usage.

5 https://mygeohub.org/tools/climatetool.