



WFDE5: bias-adjusted ERA5 reanalysis data for impact studies

Marco Cucchi¹, Graham P. Weedon², Alessandro Amici¹, Nicolas Bellouin³, Stefan Lange⁴,
Hannes Müller Schmied^{5,6}, Hans Hersbach⁷, and Carlo Buontempo⁷

¹B-Open Solutions s.r.l., Rome, Italy

²Met Office, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

³Department of Meteorology, University of Reading, Reading, Berkshire, RG6 6BB, UK

⁴Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association,
P.O. Box 60 12 03, 14412 Potsdam, Germany

⁵Institute of Physical Geography, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

⁶Senckenberg Leibniz Biodiversity and Climate Research Centre (SBiK-F),
60325 Frankfurt am Main, Germany

⁷European Centre for Medium-Range Weather Forecasts, Reading, RG2 9AX Berkshire, UK

Correspondence: Carlo Buontempo (carlo.buontempo@ecmwf.int)

Received: 10 February 2020 – Discussion started: 28 April 2020

Revised: 9 July 2020 – Accepted: 10 July 2020 – Published: 8 September 2020

Abstract. The WFDE5 dataset has been generated using the WATCH Forcing Data (WFD) methodology applied to surface meteorological variables from the ERA5 reanalysis. The WFDEI dataset had previously been generated by applying the WFD methodology to ERA-Interim. The WFDE5 is provided at 0.5° spatial resolution but has higher temporal resolution (hourly) compared to WFDEI (3-hourly). It also has higher spatial variability since it was generated by aggregation of the higher-resolution ERA5 rather than by interpolation of the lower-resolution ERA-Interim data. Evaluation against meteorological observations at 13 globally distributed FLUXNET2015 sites shows that, on average, WFDE5 has lower mean absolute error and higher correlation than WFDEI for all variables. Bias-adjusted monthly precipitation totals of WFDE5 result in more plausible global hydrological water balance components when analysed in an uncalibrated hydrological model (WaterGAP) than with the use of raw ERA5 data for model forcing.

The dataset, which can be downloaded from <https://doi.org/10.24381/cds.20d54e34> (C3S, 2020b), is distributed by the Copernicus Climate Change Service (C3S) through its Climate Data Store (CDS, C3S, 2020a) and currently spans from the start of January 1979 to the end of 2018. The dataset has been produced using a number of CDS Toolbox applications, whose source code is available with the data – allowing users to regenerate part of the dataset or apply the same approach to other data. Future updates are expected spanning from 1950 to the most recent year.

A sample of the complete dataset, which covers the whole of the year 2016, is accessible without registration to the CDS at <https://doi.org/10.21957/935p-cj60> (Cucchi et al., 2020).

Copyright statement. The works published in this journal are distributed under the Creative Commons Attribution 4.0 License. This license does not affect the Crown copyright work, which is re-usable under the Open Government Licence (OGL). The Creative Commons Attribution 4.0 License and the OGL are interoperable and do not conflict with, reduce, or limit each other.

© Crown copyright 2020

1 Introduction

The development, calibration, and evaluation of impact models require good-quality historical meteorological datasets. These are needed to both drive the impact models themselves and characterize their performances over the historical period. The availability of reliable historical runs is also critical for the preparation of impact studies using climate projections. Reanalyses have long been used for those purposes as they provide a physically consistent global reconstruction of past weather without any gap in space or time. The ERA-Interim global reanalysis for the atmosphere, land surface, and ocean waves (Dee et al., 2011) of the European Centre for Medium Range Weather Forecast (ECMWF) has been used widely as a reference by the climate community. Although reanalyses represent – by construction – the most plausible state of the atmosphere and the ocean given the observations and the forecasts from the model at a previous time-step, the coarse resolutions of models, the assumptions made in sub-grid parameterizations, and, more generally, the overall inadequacies of the modelling framework are known to induce biases with respect to ground-based observations and radiosondes. Considering that the primary goal of impact studies is to assess the climate change impacts in the real world (opposite to the modelled world), it is essential that such biases are first characterized and then, as much as practically possible, corrected for.

Recently the ERA5 reanalysis has superseded the ERA-Interim reanalysis (Hersbach et al., 2020). It is produced at ECMWF as part of the EU-funded Copernicus Climate Change Service (C3S). At the time of writing, data were available from the C3S Climate Data Store (CDS) for the period from 1979 onwards. Timely updates are provided with a 5 d latency, while a more thorough quality check is provided 2–3 months later. In 2020 the dataset will be extended back to 1950 and will then also encompass the period covered by ERA-40 (1957–2002; Uppala et al., 2005). ERA5 is based on 4D-Var data assimilation using Cycle 41r2 of the Integrated Forecasting System (IFS), which was operational at ECMWF in 2016. As such, compared to ERA-Interim (which was based on an IFS cycle that dates from 2006), ERA5 benefits from a decade of developments in model physics, core dynamics, and data assimilation. In addition to a significantly enhanced horizontal resolution (31 km grid spacing compared to 80 km for ERA-Interim), ERA5 has a number of innovative features. These include hourly

output throughout and an uncertainty estimate. The uncertainty information is obtained from a 10-member ensemble of data assimilations with 3-hourly output at half the horizontal resolution (63 km grid spacing). Compared to ERA-Interim, ERA5 also provides an enhanced number of output parameters. An overview of the main characteristics and general performance of ERA5 and a comparison with ERA-Interim is provided in Hersbach et al. (2020), while more in-depth studies of particular aspects have been reported in a growing number of publications in the scientific literature.

The move from ERA-Interim to ERA5 represents a step change in overall quality and level of detail, whose increase has been reported in a large number of publications. Several of these have been summarized in Hersbach et al. (2020), and the benefit of hourly resolution is illustrated for the December 1999 storm Lothar in that paper as well. Hersbach et al. (2019) shows the increased level in detail of precipitation over the North Atlantic.

ERA5 utilizes a vast amount of synoptic observations. The number has increased from approximately 0.75 million per day on average in 1979 to around 24 million per day by the end of 2018. Satellite radiances are the dominant and growing type of data throughout the period. The volume of conventional data has also increased steadily. In addition to observations, ERA5 relies on gridded information about radiative forcing and boundary conditions. For radiation, ERA5 includes forcings for total solar irradiance, ozone, greenhouse gases, and some aerosols developed for the World Climate Research Programme (WCRP) Coupled Model Intercomparison Project Phase 5 (CMIP5) initiative, including stratospheric sulfate aerosols. This represents a major improvement on ERA-Interim, which, for example, does not account for stratospheric sulfate aerosols due to major volcanic eruptions. Details are provided in Hersbach et al. (2015). The evolution of sea-surface temperature (SST) and sea ice cover is based on a combination of products: the UK Met Office Hadley Centre HadISST2 product for SST, the EUMETSAT OSI-SAF reprocessed product for sea ice, and the UK Met Office OSTIA product for SST and sea ice that is also used in ECMWF's operational forecasting system. Details can be found in Hirahara et al. (2016).

The EU WATCH programme produced a common framework for land surface models (LSMs) and global hydrological models (GHMs) to assess the global terrestrial hydrological cycle in the 20th and 21st centuries. This required a common meteorological forcing dataset for the 20th century, which became the WATCH Forcing Data (WFD). The WFD, based on the ERA40 reanalysis, allowed intercomparisons of hydrological models and bias correction of 21st century GCM outputs (Haddeland et al., 2011; Hagemann et al., 2011). The modelling in WATCH required sub-daily and daily average data at half-degree spatial resolution, necessitating interpolation onto the regular latitude–longitude grid, land–sea mask, and elevations used by the Climate Research Unit (CRU). The WFD methodology (Weedon et al.,

Table 1. Sources of data used to derive the WFDE5 dataset

Dataset	Summary	Location
ERA5	ECMWF reanalysis product	https://cds.climate.copernicus.eu/cdsapp#!/home (last access: 26 August 2020)
CRU TS4.03	Climate Research Unit gridded station observations (multiple variables)	http://data.ceda.ac.uk/badc/cru/data/cru_ts/cru_ts_4.03 (last access: 26 August 2020)
GPCCv2018	Global Precipitation Climatology Centre gridded station precipitation observations	https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-monthly_v2018_doi_download.html (last access: 26 August 2020)

2010, 2011) involved common processing of all terrestrial half-degree grid boxes outside Antarctica at 3-hourly steps, with elevation correction of air temperature and consequent adjustment of surface pressure, specific humidity, and downwards longwave radiation. Bias correction utilized the CRU gridded observations (New et al., 1999, 2000) of monthly average air temperature, diurnal temperature range, cloud cover (for adjusting average downwards shortwave fluxes), precipitation totals, and number of “wet” (i.e. precipitation) days. Additionally, downwards shortwave radiation was corrected for changes in multi-year tropospheric and stratospheric aerosol loading. Unlike most other reanalyses, ERA provides rainfall and snowfall rates separately, and this permitted adjustment of these rates to allow for the precipitation gauge catch corrections inherent in the observed CRU precipitation totals. Though critical for hydrological modelling, the precipitation variables are the least well constrained by surface observations, so data were provided in two versions dependent on the source of the gridded monthly observed precipitation totals – one based on CRU and the other on the “full data product” of the Global Precipitation Climatology Centre (GPCC).

Later, the WFD methodology was applied to the ERA-Interim reanalysis (Dee et al., 2011) to produce the WFDEI dataset (Weedon et al., 2014). As before, the reanalysis data were 3-hourly and interpolated onto the CRU land–sea mask. Unlike the WFD, the WFDEI includes Antarctica, and an extra processing step was introduced for the precipitation variables after correction of monthly totals and numbers of wet days and before correction of precipitation gauge biases. This involved overriding the reanalysis ratio of rainfall to snowfall in each time step in cases where the differences between the CRU grid box elevation differed substantially from ERA-Interim elevation (Weedon et al., 2014). Intermittent updates of the WFDEI beyond 2009 used the latest versions of CRU and GPCC – i.e. WFDEI files for additional years were added rather than entire new versions of the files created.

Here we describe the WFDE5 (i.e. “WATCH Forcing Data methodology applied to ERA5 reanalysis data”, C3S, 2020b), a new meteorological forcing dataset for land surface and hydrological models based on the ERA5 reanalysis (Copernicus Climate Change Service, 2017). It consists of

11 variables (see Table 2) with an hourly temporal resolution on a regular longitude–latitude half-degree grid, with global spatial coverage and values defined only for land and lake points. The dataset was derived by applying the sequential elevation and monthly bias correction methods described in Weedon et al. (2010, 2011) to half-degree aggregated ERA5 reanalysis products. The monthly observational datasets used for bias correction are CRU TS4.03 from CRU (Harris et al., 2020) for 1979 to 2018 for all variables and the GPCCv2018 full data product (Schneider et al., 2018) for rainfall and snowfall rates for 1979 to 2016. In addition, as described below, the aerosol correction step for shortwave radiation has been revised with respect to WFD and WFDEI. For an outline of the methodology applied and a reference to the observation datasets used, see Tables 1 and 2.

As a meteorological forcing dataset, WFDE5 facilitates climate impact simulations such as those carried out in the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP; Warszawski et al., 2014; Frieler et al., 2017). It can be used to directly drive historical impact simulations, which are needed for impact model validation. It can also be used as an observational reference dataset for the bias adjustment of future climate projections; these bias-adjusted climate projections can then be used to drive future climate impact projections. Both predecessors of WFDE5 have been employed for these two purposes in previous ISIMIP phases. In particular, the bias adjustment of future climate projections was done using the WFD in the ISIMIP Fast Track (Hempel et al., 2013) and the Earth2Observe, WFDEI, and ERA-Interim data merged and bias-corrected for ISIMIP (EWEMBI; Lange, 2018, 2019a) in ISIMIP2b (Frieler et al., 2017). WFDE5 will be similarly employed in the upcoming ISIMIP phase 3.

2 Dataset Processing

All computations were carried out within the CDS Toolbox, a python coding environment to retrieve, process, plot, and download data from the C3S Climate Data Store (CDS, C3S, 2020a). The CDS Toolbox scripts used to generate the dataset are publicly available at

Table 2. WFDE5 elevation and bias correction methodology outline (Weedon et al., 2010, 2011).

Variable name	Description	Units	Time step adjustments	Data used for monthly bias correction
Wind	10 m wind speed	m s^{-1}	Nil	Nil
Tair	2 m air temperature	K	Via environmental lapse rate	CRU TS4.03 temperature and diurnal temperature range
PSurf	Pressure at the surface	Pa	Via changes in Tair	Nil
Qair	2 m specific humidity	kg kg^{-1}	Via changes in Tair and PSurf	Nil
LWdown	Downward longwave radiation flux	W m^{-2}	Via fixed relative humidity and changes in Tair, PSurf, and Qair	Nil
SWdown	Downward shortwave radiation flux	W m^{-2}	Nil	CRU TS4.03 cloud cover and effects of interannual changes in atmospheric aerosol loading
Rainf _(CRU)	Rainfall rate	$\text{kg m}^{-2} \text{s}^{-1}$	Adjustment of snow/rainfall ratios	CRU TS4.03 number of wet days, CRU TS4.03 precipitation totals, ERA5 ratio of rainfall/precipitation, rainfall gauge correction
Snowf _(CRU)	Snowfall rate	$\text{kg m}^{-2} \text{s}^{-1}$	Adjustment of snow/rainfall ratios	CRU TS4.03 number of wet days, CRU TS4.03 precipitation totals, ERA5 ratio of rainfall/precipitation, snowfall gauge correction
Rainf _(CRU+GPCC)	Rainfall rate	$\text{kg m}^{-2} \text{s}^{-1}$	Adjustment of snow/rainfall ratios	CRU TS4.03 number of wet days, GPCCv2018 precipitation totals, ERA5 ratio of rainfall/precipitation, rainfall gauge correction
Snowf _(CRU+GPCC)	Snowfall rate	$\text{kg m}^{-2} \text{s}^{-1}$	Adjustment of snow/rainfall ratios	CRU TS4.03 number of wet days, GPCCv2018 precipitation totals, ERA5 ratio of rainfall/precipitation, snowfall gauge correction
ASurf	Grid-points altitude	m	Nil	Nil

Variable names and units are based on the ALMA (Assistance for Land–surface Modelling Activities) conventions (<https://www.lmd.jussieu.fr/~polcher/ALMA/>, last access: 26 August 2020). Wind, Tair, PSurf, and Qair variables have instantaneous values, while LWdown, SWdown, Rainf, and Snowf are averaged over the next hour at each date and time.

<https://doi.org/10.24381/cds.20d54e34> under a free and open licence and can be used to reproduce the dataset.

2.1 Extraction and aggregation of reanalysis data

ERA5 reanalysis data are available in the CDS on regular latitude–longitude grids at $0.25^\circ \times 0.25^\circ$, as a result of finite-element-based linear interpolation from the original reduced Gaussian grid at $\sim 0.28^\circ$, and atmospheric parameters are distributed on 37 pressure levels. They are distributed at hourly resolution, and the date and time to which each value refers to is represented using the validity date and time: for instantaneous variables, it corresponds to the date and time at which each value is considered valid; for accumulated variables, it represents the ending date and time of the

interval over which the variable is accumulated, and hence over which each value can be considered valid. Accumulation variables are aggregated over the hour ending at the validity date and time, and they are automatically converted to mean rates when retrieved from within the CDS Toolbox.

Before applying elevation and bias correction, two pre-processing steps were performed on ERA5 reanalysis data. First, in order to enable comparison and bias correction using the CRU dataset, ERA5 reanalysis were regridded to regular half-degree longitude–latitude grid, via first-order conservative remapping (Jones, 1999). Then, a backward 1 h time shift was applied to rate variables, so that values stored at each date–time represents time averages over the following hour. The latter step was taken in order to adhere to the scheme used for the WATCH Forcing Dataset (Weedon et al., 2011).

It is worth noticing that grid-points classified as belonging to land in CRU TS4.03 and GPCCv2018 datasets are not necessarily classified as land-points in ERA5 reanalysis dataset. This is especially true for coastal grid-points, for which not considering this issue often led to anomalous values in the first iteration of the WFDE5 dataset. For this reason, besides applying CRU TS4.03/GPCCv2018 to ERA5 reanalysis after half-degree regriding, an additional mask derived by ERA5 quarter-degree land–sea and lake cover mask is applied just after retrieval. In this way, the final WFDE5 dataset contains values only for grid-points that are classified as land or lake by both ERA5 and CRU.

2.2 Elevation and bias correction

Once aggregation had been performed, the sequential elevation and monthly bias correction methods of Weedon et al. (2010, 2011) were applied to the regrided data (see Table 2). The same procedures used for the creation of the WFDEI (Weedon et al., 2014) were applied, with the only exception of near-surface specific humidity (Q_{air}). For this variable, given the absence of both ERA5 near-surface specific and relative humidity from the CDS, a slightly different approach was taken: first, ERA5 vapour pressure and saturation vapour pressure at the surface, e and e_{sat} , respectively, were computed following Buck (1981). Following this, they were used to compute ERA5 relative humidity at surface as $RH = 100.0 \cdot e/e_{sat}$; finally, at this point, the algorithm described in Weedon et al. (2010) could be resumed.

Likewise for the WFD (Weedon et al., 2011) and WFDEI (Weedon et al., 2014) datasets, downward shortwave radiation was adjusted at the monthly timescale using CRU cloud cover and the local linear correlation between monthly average (aggregated) ERA5 cloud cover and downward shortwave radiation (Sheffield et al., 2006; Weedon et al., 2010).

ERA5 includes a simplified representation of the time evolution of sulfate aerosols, which interact with radiation only in that model, but otherwise does not account for the impact on surface radiative fluxes of changes in aerosol interactions with radiation (also called direct effects of aerosols) and clouds (also called first indirect effects of aerosols). To represent those impacts, aerosol corrections are calculated as monthly distributions of the anomaly in downward surface shortwave radiative flux due to aerosol–radiation and aerosol–cloud interactions over the period 1979–2018. Radiative transfer calculations, which use the tools described in Sect. 2.f.ii of Weedon et al. (2010), are based on monthly averaged distributions of tropospheric and stratospheric aerosol optical depth and cloud fraction. The time series of tropospheric optical depth for sulfate, fossil-fuel black and organic carbon, biomass burning, mineral dust, sea salt, and secondary biogenic aerosols is taken from the historical and RCP8.5 simulations by the HadGEM2-ES climate model (Bellouin et al., 2011). To correct for biases in HadGEM2-ES aerosol optical depths, these optical depths are scaled over

the whole period and for each aerosol species to match the global and monthly averages obtained by the CAMS Reanalysis of atmospheric composition (2003–2017; Inness et al., 2019), which assimilates satellite retrievals of aerosol optical depth. This bias correction was not applied in WFD and WFDEI but is now possible thanks to the availability of the CAMS Reanalysis. The time series of stratospheric aerosol optical depth is taken from the climatology by Sato et al. (1993), which has been updated to 2012 at <https://data.giss.nasa.gov/modelforce/strataer/> (last access: 26 August 2020). The years 2013–2017 are assumed to match background years, thus they replicate year 2010. That assumption is supported by the Global Space-based Stratospheric Aerosol Climatology time series (1979–2016; Thomason et al., 2018). The time series of cloud fraction is taken from CRU TS 4.03, for consistency with other aspects of the WFDE5 dataset. Surface radiative fluxes account for aerosol–radiation interactions from both tropospheric and stratospheric aerosols and for aerosol–cloud interactions from tropospheric aerosols except mineral dust. The radiative effects of aerosol–cloud interactions are assumed to scale with the radiative effects of aerosol–radiation interactions using regional scaling factors derived from HadGEM2-ES. To avoid double-counting the radiative effects of aerosol–radiation interactions by sulfate aerosols, which are to some extent already represented in ERA5, the radiative transfer calculations are repeated, this time only including sulfate aerosol–radiation interactions and the corresponding anomalies subtracted from the set of fluxes obtained previously. Atmospheric constituents other than aerosols and clouds are set to a constant standard mid-latitude summer atmosphere because their variations only have second-order effects on aerosol corrections.

Finally, similarly to the WFD and WFDEI datasets, two different WFDE5 rainfall and snowfall rates datasets, including gauge catch corrections, were generated by using either CRU TS4.03 or GPCCv2018 precipitation totals. The GPCCv2018 database includes around 3–4 times as many precipitation stations as CRU (incorporating most of the latter as a subset, Becker et al., 2013; Schneider et al., 2014) but extends only till 2016. As already pointed out in Weedon et al. (2014), during generation of the WFDE5 precipitation rates an error in the precipitation phase can arise locally where there are large elevation differences between ERA5 and CRU grids. For this reason, a further processing step was added to the WFD methodology to correct the most extreme cases of inappropriate precipitation phase: for each grid box and each calendar month over 1979–2018, records of the minimum T_{air} during rainfall and the maximum T_{air} during snowfall (“phase temperature extremes”) were stored; then, for each grid box and hourly time step, the precipitation phase was switched if the combination of the phase with the elevation and bias-corrected T_{air} were beyond a phase temperature extreme.

Elevation and bias correction was applied for all land points outside Antarctica. For grid points belonging to this

Table 3. Summary of WFDE5 dataset attributes on the C3S Climate Data Store.

Dataset attribute	Details
Horizontal coverage	Global
Horizontal resolution	$0.5^\circ \times 0.5^\circ$
Vertical coverage	Surface
Temporal coverage	– 1 January 1979 00:00:00 to 31 December 2018 23:00:00 for variables Wind, Tair, PSurf and Qair – 1 January 1979 07:00:00 to 31 December 2018 23:00:00 for variables LWdown, SWdown, Rainf _(CRU) , Snowf _(CRU) – 1 January 1979 07:00:00 to 31 December 2016 23:00:00 for variables Rainf _(CRU+GPCC) , Snowf _(CRU+GPCC)
Temporal resolution	Hourly
File format	NetCDF
Data type	Grid
Version	1.0
File naming convention	<var>_WFDE5_<reference_dataset>_<YYYYMM>_v1.0.nc, where – <var>: variable name, as in Table 2 – <reference_dataset>: one between CRU (all variables) and CRU+GPCC (Rainf and Snowf only) – <YYYYMM>: year and month

region, given the absence of observational data, only elevation correction was applied.

2.3 Higher resolution WFDE5 data

The WFDE5 has been provided at $0.5^\circ \times 0.5^\circ$ resolution rather than at $0.25^\circ \times 0.25^\circ$ as in the original ERA5 data. There are several reasons for this. The project to generate WFDE5 was also designed to deliver open-source software so that users could regenerate the data at the original or, eventually, higher resolution. Three main considerations influenced the initial generation of the WFDE5 dataset:

1. the need to generate data in time for ISIMIP3 and their reporting to the AR6 of IPCC in 2020,
2. the need to convert the existing WFDEI Fortran programs into CDS Toolbox workflows and easily test the output,
3. the requirement for appropriate and freely available global land-gridded observations for bias correction.

The first consideration meant that any procedures adopted had to be practical and fast. The simplest way to test whether the CDS Toolbox workflows were working was to apply them to ERA-Interim data and check that they correctly reproduced the WFDEI data. This implied generating output at the same resolution as the WFDEI and CRU. Additionally, ISIMIP3 only required data at $0.5^\circ \times 0.5^\circ$ since their models were set up at that resolution.

The WFDE5 CDS workflows will eventually allow users to generate higher resolution data on their own. At the moment, this can only be done using interpolated CRU TS4.03 and GPCCv2018 datasets, copies of which are hosted on

a dedicated CDS machine and made accessible through the CDS Toolbox. Another option would be to use higher-resolution observational datasets, such as quarter-degree GPCC or MSWEP (Beck et al., 2017, 2019b) for total precipitation. This option will be viable once additional datasets can be hosted on the C3S Climate Data Store.

3 Data and software access

The WFDE5 dataset is distributed by the Copernicus Climate Change Service (C3S) through its Climate Data Store (CDS) as monthly files in NetCDF format and can be downloaded at <https://doi.org/10.24381/cds.20d54e34> (C3S, 2020b). It uses a full half-degree grid (720×360 grid boxes) with the sea and large lakes flagged as missing data, comprising a total of 92889 land points (Antarctica included). General dataset attributes are described in Table 3. A sample of the complete dataset, which covers the whole of the year 2016, is accessible without registration to the CDS at <https://doi.org/10.21957/935p-cj60> (Cucchi et al., 2020).

All the CDS Toolbox workflows used to generate WFDE5 are publicly available (<https://doi.org/10.24381/cds.20d54e34>) and can be used to regenerate samples of the dataset. Furthermore, as ERA5 progresses, using these applications it will be possible to expand WFDE5 dataset back to the start of 1950 and forward beyond 2018.

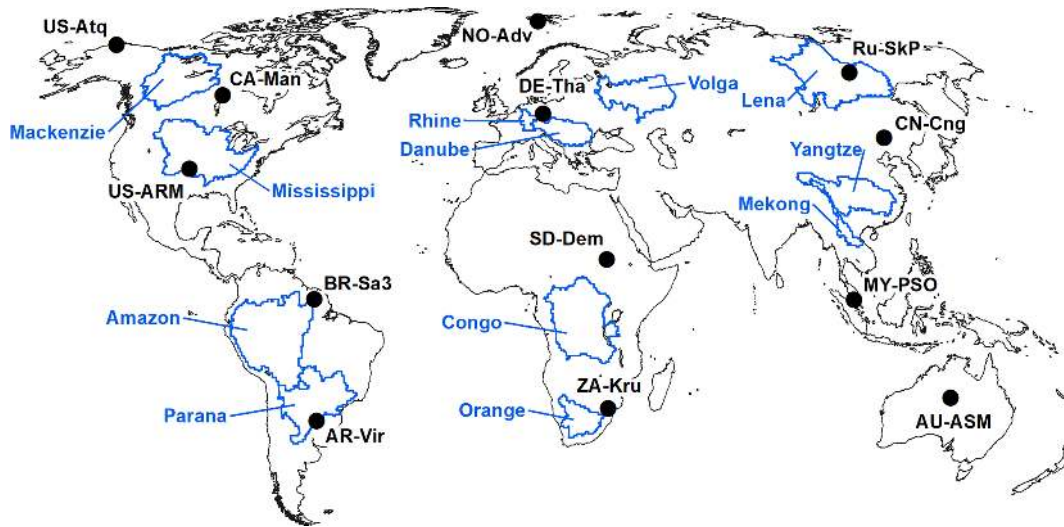


Figure 1. Location of FLUXNET2015 sites used to evaluate ERA5, WFDE5, and WFDEI, as well as river basin outlines for the hydrological assessment.

4 Evaluation

4.1 Previous analyses

Beck et al. (2019a) assessed multiple precipitation datasets at daily time steps against radar and precipitation gauge observations across the contiguous USA. Their analysis included ERA5, ERA-Interim and WFDEI precipitation adjusted to GPCP totals. They demonstrated that against observations ERA5 precipitation provides a significant improvement over both ERA-Interim and WFDEI precipitation. Albergel et al. (2018) used the ISBA LSM to assess the use of ERA5 versus ERA-Interim forcing. They assessed performance against a wide variety of observed hydrological and vegetation-related variables. Significant improvements were demonstrated in the simulation of the hydrological cycle using ERA5, which they mostly attributed to better precipitation. There were small changes related to vegetation modelling. For a region with a low density of gauges in Iran, Fallah et al. (2020) showed that ERA5 precipitation is closer to local observations than ERA-Interim but that GPCPv2018 (used here in bias correction or ERA5) is substantially better.

4.2 Comparison with FLUXNET2015 and WFDEI

The FLUXNET2015 (FN2015) meteorological data (Chu, 2015; Pastorello et al., 2017) are not included in the data assimilation of the ERA5 reanalysis. Therefore, these data provide an opportunity to assess the degree to which the ERA5 and WFDE5 meteorological variables agree with surface observations. Despite there being over 200 FN2015 sites globally, they are highly clustered within Europe and North America. In order to provide a fairly uniform global assessment, 13 sites with at least 3 years of data have been selected from 12 countries spanning a wide range of longitudes

and latitudes (Fig. 1, Table 4). The primary purpose of the FN2015 meteorological dataset is to provide data for forcing LSMs to allow comparison with the FN2015 surface exchange fluxes of energy and carbon. As such, the FN2015 meteorological variables have been gap-filled using ERA-Interim data to allow modelling without missing data. To avoid biasing the comparisons made here, only meteorological values that are measurements have been used (i.e. at times and locations where the FN2015 tier 1 quality flag is 0). Unfortunately, this means that some FN2015 sites do not provide observations for some variables at any time step (“missing variables” in Table 4).

Two pairs of comparisons have been made: firstly for ERA5 (aggregated to half degree) versus FN2015, as well as for WFDE5 versus FN2015 at an hourly time step. This required converting the half-hourly FN2015 data to hourly steps and aligning the time stamps since ERA5 is based on UTC instead of local time. ERA5 does not provide specific humidity so Q_{air} was calculated using the 2 m air temperature, surface pressure, and relative humidity using equations 4 and 6 of Buck (1981). The second comparisons were for WFDEI versus FN2015 and for WFDE5 versus FN2015 at 3-hourly time steps, again with alignment of time stamps. At each site mean bias error (MBE), mean absolute error (MAE), and correlation were calculated. MAE was used instead of root-mean-square error since the former provides a less ambiguous basis for assessment (Willmott and Matsuura, 2005). Since the data are time series there is considerable serial correlation leading to spuriously high values. Consequently, the correlations of the previously pre-whitened time series – i.e. adjusted to remove lag-1 autocorrelation – are also reported as “adjusted correlation” (Ebisuzaki, 1997). Data for individual sites are reported for ERA5 and WFDE5 versus FN2015 (hourly) in Tables A1–A8 and for WFDEI

Table 4. Selected FLUXNET2015 sites.

Site code	Country	Site name	Longitude	Latitude	Start year	End year	Missing variables	DOI
US-Atq	USA	Atqauk	157.41° W	70.47° N	2006	2008	LWdown	https://doi.org/10.18140/FLX/1440067
CA-Man	Canada	Manitoba	98.48° W	55.88° N	2006	2008	PSurf, LWdown	https://doi.org/10.18140/FLX/1440035
US-ARM	USA	ARM Southern Great Plain site – Lamont	97.49° W	36.61° N	2010	2012		https://doi.org/10.18140/FLX/1440066
BR-Sa3	Brazil	Santarem km67 primary forest	54.97° W	3.02° N	2001	2003		https://doi.org/10.18140/FLX/1440033
AR-Vir	Argentina	Virasoro	56.19° W	28.24° S	2010	2012	LWdown, Precip	https://doi.org/10.18140/FLX/1440192
NO-Adv*	Norway	Adventdalen	15.92° E	78.19° N	2012	2014		https://doi.org/10.18140/FLX/1440241
DE-Tha	Germany	Tharandt	13.57° E	50.96° N	2012	2014		https://doi.org/10.18140/FLX/1440152
SD-Dem	Sudan	Demokeya	30.48° E	13.28° N	2007	2009	LWdown	https://doi.org/10.18140/FLX/1440186
ZA-Kru	South Africa	Skukuza	31.50° E	25.02° S	2008	2010	PSurf, LWdown	https://doi.org/10.18140/FLX/1440188
RU-SkP	Russia	Yakutsk Spasskaya Pad larch	129.17° E	62.26° N	2010	2014	Precip	https://doi.org/10.18140/FLX/1440243
CN-Cng	China	Chanling	123.51° E	44.59° N	2008	2010		https://doi.org/10.18140/FLX/1440209
MY-PSO	Malaysia	Pasoh Forest Reserve	102.31° E	2.97° N	2007	2009	PSurf	https://doi.org/10.18140/FLX/1440240
AU-ASM	Australia	Alice Springs	133.25° E	22.28° S	2011	2013		https://doi.org/10.18140/FLX/1440194

* NO-Adv is now designated as SL-Adv (i.e. within Svalbard). Precip stands for precipitation. Note that “missing variables” refers to tier 1 items provided by FLUXNET2015 as entirely gap-filled, i.e. not measured, values.

and WFDE5 versus FN2015 (3-hourly) in Tables A9–A16. Average metrics for the pairs of comparisons are shown in Fig. 2 and Tables A17 and A18.

At hourly steps on average there are no significant differences in MBE, MAE, correlation or adjusted correlation between ERA5 versus FN2015 and WFDE5 versus FN2015, for all variables apart from two (Fig. 2a). For air temperature the MBE is slightly better (closer to zero) for WFDE5, whereas the MAE is slightly worse (larger) for WFDE5. On the other hand, for specific humidity both the MBE and MAE are slightly worse for WFDE5. These results indicate that the bias and elevation corrections incorporated into the WFDE5 have had little overall effect on the performance against surface observations compared to ERA5.

At three hourly steps, for all variables apart from precipitation, the average MBE plus 95 % confidence intervals overlap zero for WFDEI and WFDE5 (Fig. 2b). For wind speed, downwards longwave and downwards shortwave the MAE is slightly better (smaller) for WFDE5 than WFDEI. For all variables aside from precipitation, the MAE, correlation, and adjusted correlation are slightly better for WFDE5 than WFDEI. For precipitation the MBE is slightly better and the correlation slightly higher for WFDE5 versus WFDEI when corrected using the GPCC-corrected, rather than CRU-corrected, precipitation totals. These results indicate that, on average, at the FN2015 sites selected, WFDE5 performs better than WFDEI against the observations. Note that the average results in Fig. 2b and Table A17 hide the fact that for all metrics WFDEI data provide better results (MBE closer to 0.0, MAE lower, correlation higher) for some individual sites than WFDE5 (Tables A9–A16). On the other hand, for wind (speed) and precipitation (CRU- and GPCC-corrected)

the correlation and adjusted correlation are better for WFDE5 than WFDEI at every site.

Both WFDEI and WFDE5 in 2017 and 2018 are corrected using CRU TS4.03, so at monthly scales and longer there will be only small differences. However, at sub-monthly timescales, aside from advances in the processing system between the reanalyses used, it is likely that the better performance of WFDE5 is linked to superior spatial variability of ERA5 (data aggregated for WFDE5) versus ERA-Interim (data interpolated for WFDEI). This can be seen in the higher-resolution features of daily average temperature for a single day in January 2018 in North America and Central America in the WFDE5 data (Fig. 3).

4.3 Validation with a global hydrological model

Of great importance for driving impact models such as global hydrological models is the climate forcing input, since the water balance components are highly dependent on it (Müller Schmied et al., 2016). In order to test WFDE5 in terms of suitability for use with an impact model, the global water availability and water use model WaterGAP (version 2.2c, Müller Schmied et al., 2016) was used. WaterGAP calculates water storages and fluxes on global land area (except Antarctica) on a $0.5^\circ \times 0.5^\circ$ resolution (55×55 km at the Equator) and incorporates human interventions such as human water use and man-made reservoirs. Forcing requirements are daily values for precipitation (sum of rainfall and snowfall), average temperature, downwards shortwave radiation, and downwards longwave radiation. For model-specific details, the reader is referred to Müller Schmied et al. (2016, 2014), and Döll et al. (2003). Despite the possibility of cal-

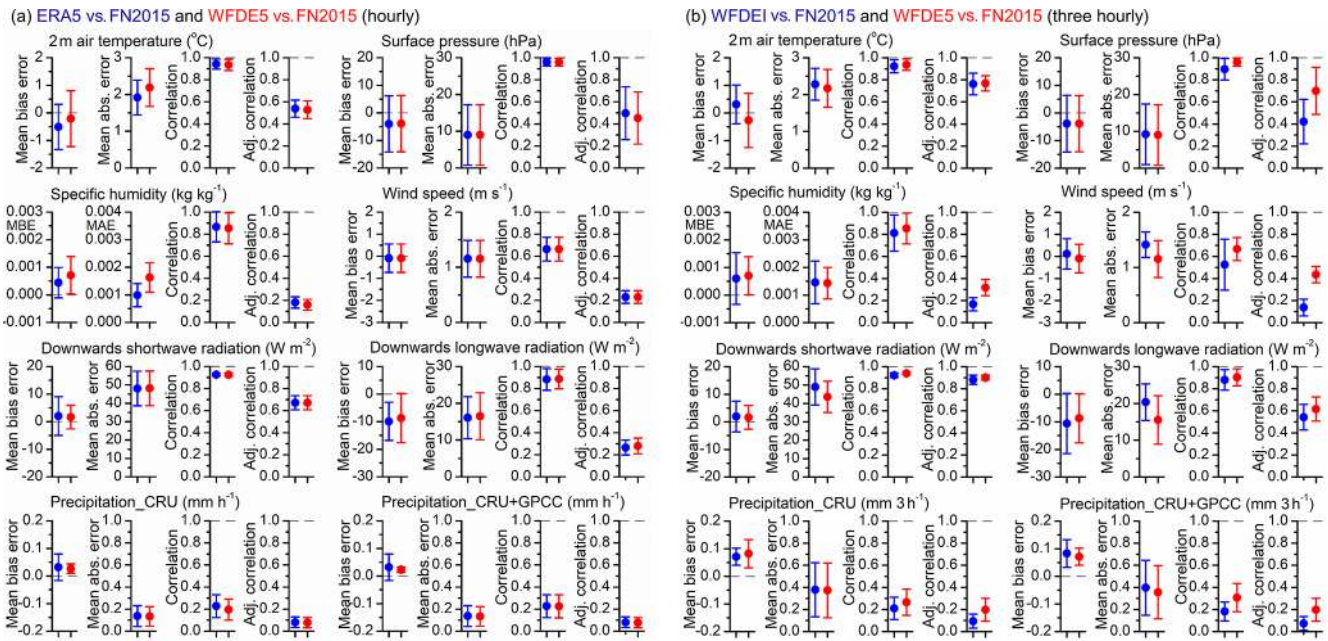


Figure 2. Average metrics (mean \pm 95 % confidence interval of the mean) for (a) ERA5 versus FN2015 (blue) and WFDE5 versus FN2015 (red) at hourly time steps (see Table A17); (b) WFDEI versus FN2015 (blue) and WFDE5 versus FN2015 (red) at three-hourly time steps (see Table A18).

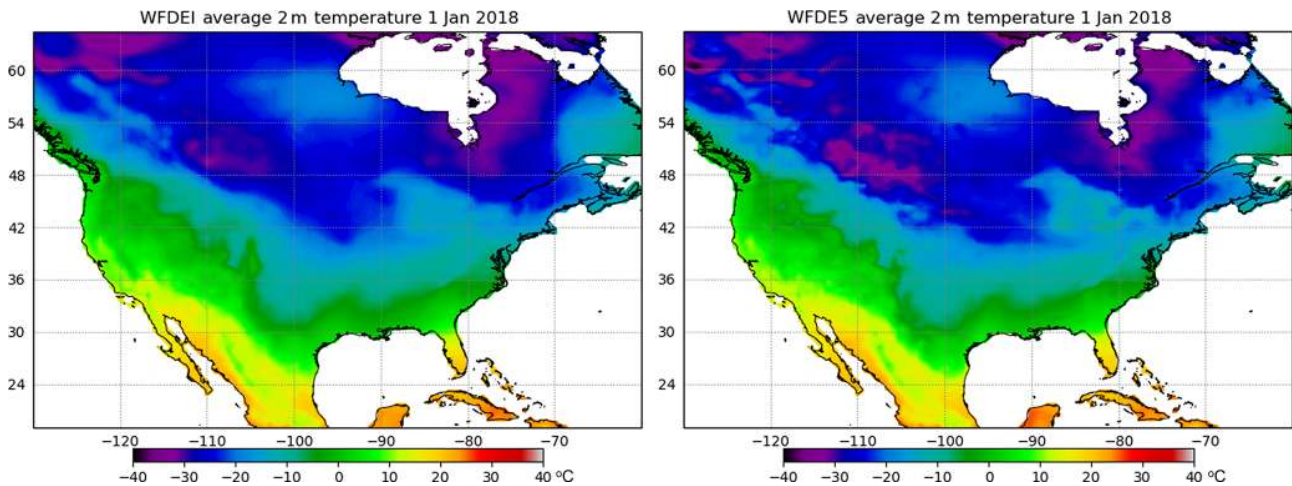


Figure 3. Average 2 m temperature on 1 January 2018 for north and central America.

ibrating the model, WaterGAP was run with an uncalibrated setup (model parameter γ set to 2, whereas CFA and CFS are set to 1 globally; details can be found in Müller Schmied et al., 2014). This parameter choice was designated to mimic the behaviour in a typical impact model and also due to time and technical constraints (a time series start year of 1920 or earlier is required for standard calibration). The model was driven by ERA5, WFDE5, and WFDEI (the latter two with monthly precipitation scaled to both GPCC and CRU) and was assessed in terms of resulting water balance components

(Table 5), model efficiency (Fig. 4), and river discharge seasonality for selected large river basins (Fig. 5).

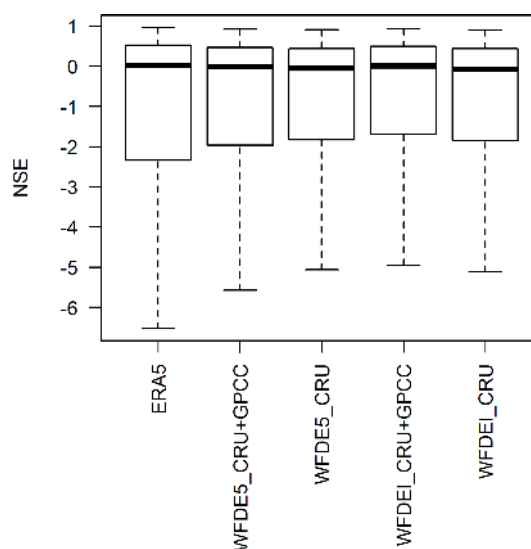
The long-term-annual water balance shows considerably (around 10 %) higher precipitation (P) for ERA5 compared to the WFDE5 adjustments to GPCC or CRU which results in higher values for actual evapotranspiration (AET) and greater river discharge to oceans and inland sinks (Q) (Table 5). The general reduction of mean global precipitation from 120 000 km³ yr⁻¹ for ERA5 to observation-based datasets with around 111 000 km³ yr⁻¹ for WFDE5 is consistent with previous estimates (109 631

Table 5. Long-term-annual water balance components [$\text{km}^3 \text{yr}^{-1}$] as simulated with uncalibrated WaterGAP 2.2c for 1981–2010 and for global land area (except Antarctica and Greenland).

No.	Component	ERA5	WFDE5_CRU		WFDEI_CRU	
			+ GPCC	WFDE5_CRU	+ GPCC	WFDEI_CRU
1	Precipitation	120 245	111 529	110 981	111 616	111 554
2	Actual evapotranspiration	76 695	73 430	74 702	73 540	74 230
3	River discharge to oceans and inland sinks	43 623	38 135	36 310	38 088	37 320
4	Total (actual) water consumptions (rows 5 + 6)	1105	1183	1151	1103	1086
5	Net (actual) abstraction from surface water	1241	1359	1318	1246	1223
6	Net abstraction from groundwater	−136	−176	−167	−143	−137
7	Change of total water storage	−74	−36	−31	−12	4
8	Long-term annual water balance error	0.16	0.15	0.14	0.14	0.14

to $111\,050 \text{ km}^3 \text{ yr}^{-1}$ for the time span 1971–2000 for the snow-undercatch-corrected climate forcings in Tables 3 and 4 of Müller Schmied et al., 2016). Even though WaterGAP was not calibrated, AET and Q are well within the estimates of other models or datasets (AET $62\,800$ – $75\,981 \text{ km}^3 \text{ yr}^{-1}$ and Q $34\,400$ – $44\,560$ for most assessments according to Müller Schmied et al., 2014, Table 5). The differences in GPCC and CRU dataset versions to adjust ERA5 (ERA-Interim) precipitation for WFDE5 (WFDEI) is substantially smaller for GPCC (precipitation difference: $87 \text{ km}^3 \text{ yr}^{-1}$ for WFDE5_CRU+GPCC versus WFDEI_CRU+GPCC) compared to CRU (precipitation difference: $573 \text{ km}^3 \text{ yr}^{-1}$ for WFDE5_CRU versus WFDEI_CRU). Consequently, differences in simulated river discharge are higher for WFDE5_CRU versus WFDEI_CRU ($1010 \text{ km}^3 \text{ yr}^{-1}$) compared to WFDE5_CRU+GPCC versus WFDEI_CRU+GPCC ($47 \text{ km}^3 \text{ yr}^{-1}$). This implies that the choice of precipitation bias adjustment target (CRU or GPCC) impacts water balance components.

The performance of the uncalibrated model runs have been assessed using the widely used Nash–Sutcliffe efficiency metric (NSE, Nash and Sutcliffe, 1970) relative to monthly time series of GRDC station observed discharge. A total of 1216 stations have been used out of the usual 1319 stations used for WaterGAP calibration (Müller Schmied et al., 2014) constrained by data availability for at least 1 year in the time span of the forcing. The optimum NSE is 1, and the value can become infinitely negative, but below 0 the simulation is not better than the average of the observations (Nash and Sutcliffe, 1970). The median performances of the model runs are similar and around the value 0, with some ranging towards optimum but also towards negative NSE values. Note that around 16 % to 17 % of the stations are consistently outside of the limits of the boxplots ($\text{NSE} > 1.5 \times \text{inter quartile range}$) towards negative values and not displayed. Generally, the variants scaled to GPCC tend to have a slightly better performance than the values scaled to CRU. Typically, the performance increases as a result of calibrating the model (see Müller Schmied et al., 2014, Fig. 6), so the NSE values

**Figure 4.** Model efficiency for the uncalibrated runs of the climate forcings in this assessment using monthly time series of 1216 GRDC stations.

reported here should not be wrongly interpreted as the result of the poor quality of the forcing data but more in the sense of that uncalibrated impact models could reach – in principle – similar efficiencies independently of the forcing data assessed here (with slight advantages of the bias-adjusted WFDE5 data compared to direct use of ERA5, Fig. 4).

In Fig. 5 discharge seasonality is shown with GRDC observations in black; see Fig. 1 for basin outlines. The figure shows the effect of adjusting precipitation from ERA5 (red in Fig. 5). For most basins, but not for all (e.g. Mississippi), the adjustment to CRU or GPCC-precipitation leads to a reduction of river discharge; this is substantial for some basins, e.g. Yangtze and Amazon. This does not necessarily lead to a better agreement with the observations (e.g. Amazon, Mackenzie, Lena), but for a number of basins it does (e.g. Congo, Orange, Mekong, Danube). Interestingly, the ef-

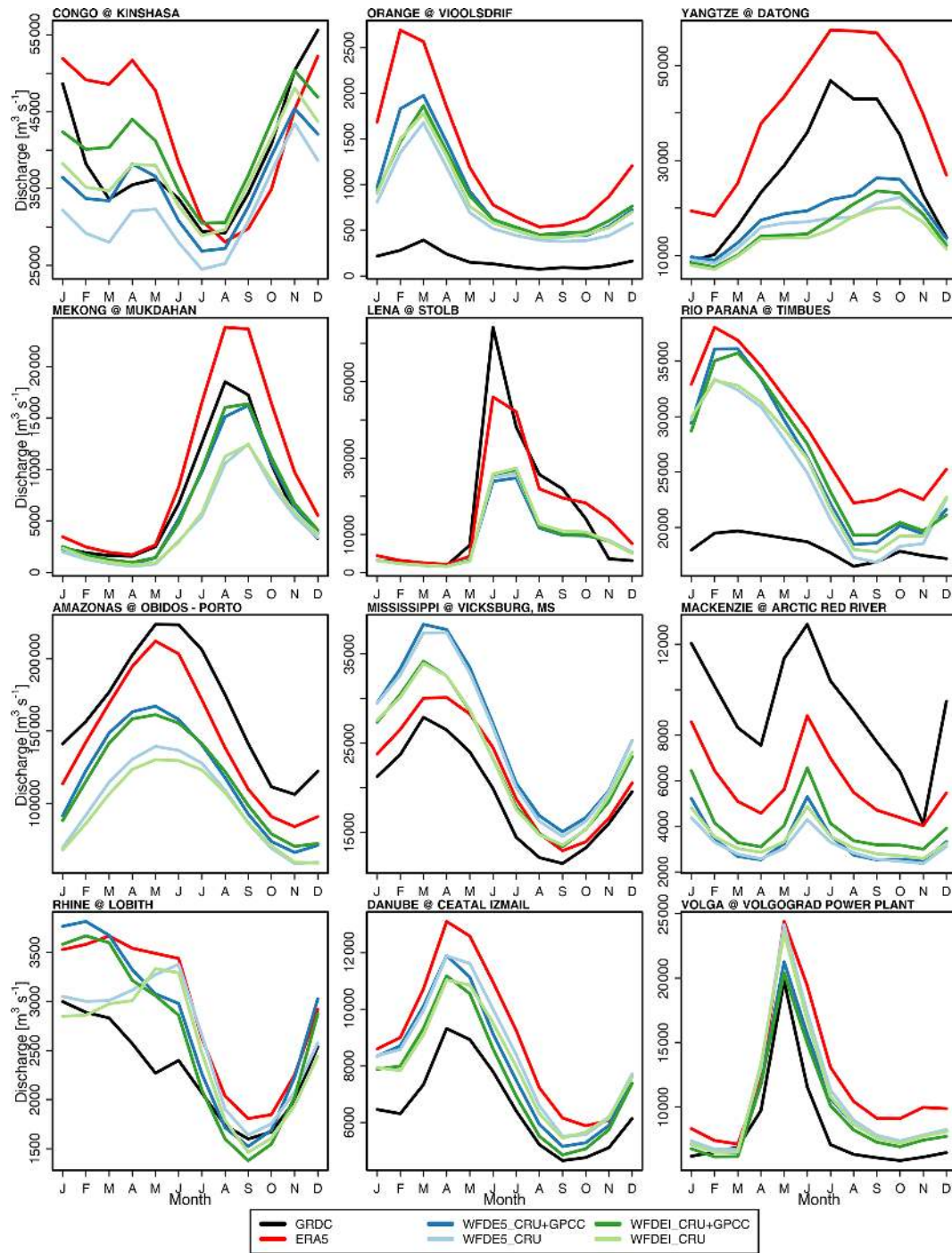


Figure 5. Seasonality of observed river discharge and uncalibrated WaterGAP runs for selected large river basins (Fig. 1).

fect of the dataset chosen to adjust precipitation (CRU versus GPCC) is important for some basins (e.g. Mekong, Amazon). However, this is not relevant for other basins (e.g. Mississippi, Danube) where differences in WFDE5 and WFDEI compared to ERA5 and ERA-Interim for variables other than precipitation lead to different discharge simulations. An overview of spatial differences in long-term average precipi-

tation between WFDE5 and ERA5 and can be found in Fig. 6 and help to interpret the patterns observed in Fig. 5. Spatial differences of the other variables used for WaterGAP are shown in Figs. A1–A3.

The validation with WaterGAP showed that using WFDE5 generally results in similar results to using WFDEI and should be preferred to using ERA5 directly. Nevertheless,

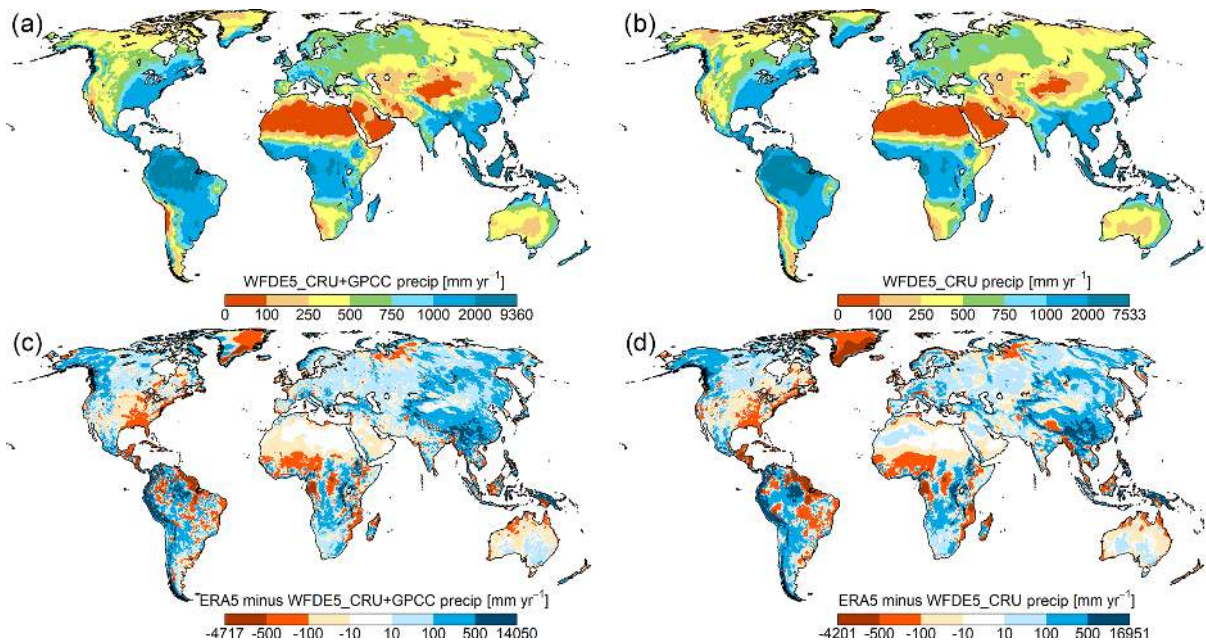


Figure 6. Long-term (1979–2016) average precipitation of the climate forcings, displayed as absolute number for WFDE5_CRU+GPCC (a), WFDE5_CRU (b), and differences to ERA5, computed as ERA5 minus WFDE5_CRU+GPCC (c) and ERA5 minus WFDE5_CRU (d). All values are given in millimetres per year.

this assessment was done using uncalibrated runs, thus a proper calibration to discharge observations could highlight the full benefit of WFDE5 compared to ERA5, but this is outside of the scope of this paper.

5 Application in ISIMIP

The WFDE5 dataset will be employed to drive historical impact simulations and bias-adjust future climate projections in the upcoming ISIMIP phase 3. The dataset is well suited for these purposes in particular thanks to its inter-variable consistency, which matters for the simulation of extreme climate impact events (Zscheischler et al., 2019). Thanks to a new bias adjustment method that is applied in ISIMIP phase 3, that is able to adjust inter-variable statistical dependencies (Lange, 2019b, 2020), the inter-variable consistency of WFDE5 will be beneficial for the bias adjustment of future climate projections as well.

Instead of using WFDE5 directly for these purposes, a derived dataset covering land and ocean with daily temporal resolution and including additional variables will be used in ISIMIP phase 3. This derived dataset consists of WFDE5 over land merged with ERA5 over the ocean (W5E5; Lange, 2019b). It covers land and ocean to facilitate impact studies everywhere and prevent mismatches between land–sea masks used by impact models. It has daily temporal resolution because that is sufficient to drive most impact models taking part in ISIMIP. Additional variables (2 m relative humidity, sea level pressure, total precipitation, daily maxi-

mum 2 m air temperature, daily minimum 2 m air temperature) derived from those included in WFDE5 are included in W5E5 to meet additional impact model requirements. More information about the W5E5 dataset is provided by Lange (2019b).

6 Code and data availability

The full information regarding access to the datasets can be found in Sect. 3.

7 Conclusions

The WFDE5 dataset will be useful for forcing surface models and especially for near-recent hydrological and agricultural analyses. It will also be used for bias correction of the CMIP6 GCM model output in the third phase of ISIMIP. WFDE5 benefits from the improvements of ERA5 compared to ERA-Interim, as well as from the additional corrections of temperature, precipitation, and shortwave radiation described above.

WFDE5 is provided at hourly time steps versus 3-hourly time steps for WFDEI. Comparison to observations from 13 FLUXNET2015 sites distributed globally shows that, on average, WFDE5 is superior to WFDEI for all variables in terms of mean absolute error and correlation. For precipitation and wind speed WFDE5 is superior to WFDEI at all 13 sites. Although both datasets are provided at 0.5° resolution, WFDE5 has a greater spatial variability (Fig. 3) since it is ob-

tained by aggregation of higher-resolution ERA5 data rather than by interpolation of lower-resolution ERA-Interim data used in WFDEI. Initial analysis using an uncalibrated hydrological model (WaterGAP) has demonstrated that the bias correction to CRU or GPCC precipitation totals results in lower discharge throughout the year bringing the global hydrological balance into better agreement with previous studies. The corrections result in improvements towards observations relative to the use of unaltered ERA5 forcing (e.g. in the Congo, Orange, and Danube basins).

Currently the WFDE5 dataset spans from the start of 1979 to the end of 2018 (end of 2016 for Rainf_WFDE5_CRU+GPCC, and Snowf_WFDE5_CRU+GPCC). However, the open source Python code within the Climate Change Service Toolbox will allow users to expand the coverage back to the start of 1950 and forwards through 2019 and later for themselves. The data have been created at 0.5° resolution to match the CRU grid, but gridded observations of precipitation totals are already available from GPCC at 0.25° and MSWEPv2 at 0.1° (Beck et al., 2019b). The future availability of gridded observations of near-surface temperature, diurnal temperature range, cloud cover, aerosol loading, and numbers of wet days would allow creation of WFDE5 data at higher spatial resolution than the current dataset.

Appendix A

Table A1. Tair metrics for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

Tair (°C)	No. of points	Mean	MBE		MAE		Correlation		Adj. correlation	
			Fluxnet 2015	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5
US-Atq	26 227	-10.127	0.529	0.514	2.035	2.712	0.984	0.977	0.451	0.401
CA-Man	20 062	0.741	-1.620	-2.138	1.949	2.452	0.992	0.988	0.727	0.712
US-ARM	23 442	15.451	0.856	1.999	1.703	1.999	0.978	0.977	0.745	0.744
BR-Sa3	24 500	25.874	0.449	1.125	1.668	2.044	0.724	0.689	0.381	0.383
AR-Vir	18 209	21.750	-0.199	0.034	1.501	1.541	0.958	0.958	0.574	0.572
NO-Adv	12 039	-2.298	-1.797	-1.657	2.080	2.093	0.980	0.969	0.469	0.409
DE-Tha	26 298	9.324	-1.060	-0.023	1.588	1.714	0.973	0.968	0.634	0.629
SD-Dem	24 624	27.058	0.072	1.217	1.386	1.858	0.956	0.952	0.638	0.640
ZA-Kru	24 825	21.701	-1.424	-0.631	2.194	2.446	0.887	0.878	0.425	0.427
RU-SkP	22 123	-2.740	-3.891	-4.262	4.276	4.650	0.984	0.982	0.331	0.330
CN-Cng	24 117	6.760	0.321	0.303	1.320	1.354	0.994	0.994	0.526	0.520
MY-PSO	26 295	25.066	0.753	1.153	1.155	1.401	0.884	0.879	0.593	0.594
AU-ASM	26 239	22.695	-0.269	-0.353	2.082	2.200	0.950	0.946	0.502	0.501

Table A2. PSurf metrics for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

PSurf (hPa)	No. of points	Mean	MBE		MAE		Correlation		Adj. correlation	
			Fluxnet 2015	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5
US-Atq	17 345	1012.4	-2.8	-2.8	2.8	2.8	0.986	0.986	0.050	0.050
US-ARM	24 196	976.3	0.4	0.5	0.6	0.6	0.997	0.996	0.733	0.656
BR-Sa3	22 274	986.1	15.0	15.0	15.0	15.0	0.880	0.880	0.051	0.060
AR-Vir	18 048	999.2	4.1	4.1	4.1	4.1	0.992	0.992	0.883	0.797
NO-Adv	12 038	1011.3	-39.5	-39.4	39.5	39.4	0.994	0.994	0.847	0.845
DE-Tha	26 301	972.3	-9.1	-8.9	9.1	8.9	0.997	0.997	0.378	0.166
SD-Dem	6636	949.6	-8.4	-8.2	8.5	8.2	0.978	0.972	0.551	0.695
RU-SkP	22 128	987.5	-2.3	-2.4	2.3	2.4	0.996	0.996	0.913	0.726
CN-Cng	24 117	992.0	5.0	4.9	5.6	5.6	0.867	0.867	0.312	0.085
AU-ASM	26 200	945.0	-2.1	-2.2	2.9	2.9	0.930	0.930	0.254	0.453

Table A3. Qair metrics for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

Qair (kg kg ⁻¹)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
US-Atq	26 304	0.00242	-0.00010	-0.00031	0.00031	0.00046	0.963	0.946	0.257	0.169
CA-Man	26 235	0.00410	-0.00049	-0.00063	0.00088	0.00095	0.928	0.925	0.135	0.158
US-ARM	26 294	0.00771	0.00026	0.00053	0.00095	0.00106	0.955	0.949	0.174	0.172
BR-Sa3	26 280	0.01537	0.00288	0.00365	0.00295	0.00370	0.208	0.201	-0.020	-0.033
AR-Vir	26 299	0.01123	0.00056	0.00074	0.00130	0.00142	0.866	0.859	0.120	0.126
NO-Adv	26 234	0.00271	-0.00033	-0.00035	0.00039	0.00043	0.979	0.964	0.228	0.199
DE-Tha	26 304	0.00573	-0.00030	0.00077	0.00049	0.00086	0.975	0.969	0.291	0.268
SD-Dem	26 304	0.00800	0.00042	0.00089	0.00106	0.00134	0.963	0.959	0.205	0.192
ZA-Kru	26 304	0.01072	0.00016	0.00079	0.00095	0.00142	0.922	0.894	0.169	0.074
RU-SkP	21 505	0.00358	0.00007	0.00004	0.00054	0.00058	0.970	0.966	0.128	0.110
CN-Cng	26 303	0.00486	0.00023	0.00021	0.00052	0.00051	0.988	0.987	0.312	0.301
MY-PSO	26 304	0.01634	0.00157	0.00201	0.00160	0.00203	0.577	0.520	0.132	0.133
AU-ASM	26 304	0.00606	0.00082	0.00083	0.00098	0.00103	0.962	0.853	0.218	0.207

Table A4. Wind metrics for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

Wind (ms ⁻¹)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
US-Atq	19 078	3.587	1.233	1.233	1.454	1.454	0.840	0.840	0.251	0.251
CA-Man	17 111	3.403	-0.251	-0.251	0.684	0.684	0.790	0.790	0.254	0.254
US-ARM	21 285	4.636	-0.320	-0.320	1.103	1.103	0.816	0.816	0.308	0.308
BR-Sa3	23 365	2.229	-0.542	-0.542	0.835	0.835	0.296	0.296	0.113	0.113
AR-Vir	18 002	2.317	0.821	0.821	0.998	0.998	0.677	0.677	0.217	0.217
NO-Adv	6 778	5.394	-2.690	-2.690	2.744	2.744	0.714	0.714	0.354	0.354
DE-Tha	25 974	3.033	-0.157	-0.157	0.899	0.899	0.672	0.672	0.154	0.154
SD-Dem	24 620	2.848	0.897	0.897	1.274	1.273	0.583	0.583	0.271	0.271
ZA-Kru	24 825	3.242	-1.160	-1.160	1.309	1.309	0.645	0.645	0.241	0.241
RU-SkP	16 790	2.748	-0.042	-0.042	0.616	0.616	0.786	0.786	0.182	0.182
CN-Cng	24 096	3.670	0.080	0.080	0.904	0.904	0.821	0.821	0.305	0.305
MY-PSO	26 061	1.785	-0.303	-0.303	0.769	0.769	0.317	0.317	0.019	0.019
AU-ASM	26 213	2.565	1.289	1.289	1.430	1.430	0.675	0.675	0.322	0.322

Table A5. SWdown metrics for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

SWdown (W m ⁻²)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
US-Atq	25 819	106.825	-16.268	-16.318	31.256	31.304	0.932	0.932	0.679	0.679
CA-Man	22 867	132.639	0.883	0.698	42.677	42.601	0.916	0.916	0.612	0.612
US-ARM	24 056	195.043	1.182	2.389	46.446	46.531	0.951	0.951	0.785	0.784
BR-Sa3	24 000	186.600	8.492	8.470	65.037	65.200	0.892	0.892	0.568	0.568
AR-Vir	15 829	93.032	3.217	3.742	33.260	33.490	0.925	0.924	0.654	0.654
NO-Adv	12 174	87.650	4.177	4.280	28.848	28.880	0.901	0.900	0.477	0.477
DE-Tha	26 164	124.441	1.696	3.955	39.427	39.607	0.927	0.927	0.601	0.601
SD-Dem	24 386	257.052	12.876	12.149	50.778	50.612	0.965	0.965	0.878	0.878
ZA-Kru	20 650	196.204	2.050	1.279	53.103	52.882	0.932	0.932	0.715	0.715
RU-SkP	21 726	128.763	1.521	0.161	45.401	45.304	0.916	0.916	0.631	0.631
CN-Cng	25 305	163.944	9.518	8.084	41.196	41.016	0.949	0.949	0.765	0.765
MY-PSO	26 084	193.032	0.169	-5.169	62.153	62.170	0.910	0.909	0.614	0.613
AU-ASM	26 210	255.772	-3.263	-2.342	84.963	86.203	0.918	0.917	0.740	0.740

Table A6. LWdown metrics for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

LWdown (W m ⁻²)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
US-ARM	24 106	335.275	-5.157	-1.555	13.372	13.281	0.964	0.962	0.342	0.366
BR-Sa3	18 705	417.810	-3.479	2.234	10.979	10.591	0.699	0.729	0.303	0.336
NO-Adv	11 092	287.082	-26.027	-28.481	30.119	32.124	0.880	0.868	0.153	0.159
DE-Tha	26 301	315.195	-7.511	-2.397	15.473	14.744	0.901	0.898	0.208	0.216
RU-SkP	21 963	261.041	-17.763	-20.020	22.163	23.999	0.972	0.970	0.169	0.182
CN-Cng	21 700	287.760	-11.902	-11.579	14.911	14.822	0.980	0.980	0.318	0.325
MY-PSO	26 263	417.346	-2.314	-1.719	11.217	10.728	0.700	0.722	0.256	0.277
AU-ASM	26 221	346.663	-5.025	-6.203	10.320	11.483	0.969	0.964	0.372	0.377

Table A7. Precipitation (Rainf + Snowf) metrics, corrected using CRU totals, for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

<i>P</i> CRU (mm h ⁻¹)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
US-Atq	26 133	0.104	0.018	0.014	0.035	0.032	0.135	0.066	0.046	0.022
CA-Man	13 634	0.042	0.047	0.029	0.097	0.085	0.232	0.195	0.079	0.074
US-ARM	24 114	0.056	0.024	0.045	0.098	0.114	0.292	0.272	0.112	0.102
Br-Sa3	26 280	0.161	0.089	0.078	0.374	0.366	0.044	0.038	0.014	0.013
NO-Adv	6516	0.018	0.048	0.052	0.064	0.071	0.366	0.239	0.097	0.060
DE-Tha	26 304	0.097	0.000	-0.013	0.113	0.104	0.412	0.412	0.180	0.173
SD-Dem	24 621	0.032	-0.008	0.008	0.053	0.068	0.061	0.060	0.006	0.002
ZA-Kru	24 818	0.044	0.043	0.033	0.110	0.101	0.179	0.174	0.045	0.038
CN-Cng	24 117	0.037	0.023	0.017	0.062	0.059	0.516	0.455	0.233	0.194
MY-PSO	26 301	0.220	0.068	0.029	0.455	0.420	0.079	0.075	0.017	0.014
AU-ASM	26 234	0.032	0.002	0.001	0.051	0.055	0.182	0.159	0.064	0.072

Table A8. Precipitation (Rainf + Snowf) metrics, corrected using GPCC totals, for each FLUXNET2015 site: ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps.

<i>P</i> GPCC (mm h ⁻¹)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
US-Atq	26 133	0.104	0.018	0.006	0.035	0.024	0.135	0.098	0.046	0.025
CA-Man	13 634	0.042	0.047	0.042	0.097	0.094	0.232	0.220	0.079	0.079
US-ARM	24 114	0.056	0.024	0.027	0.098	0.099	0.292	0.312	0.112	0.115
BR-Sa3	26 280	0.161	0.089	0.045	0.374	0.333	0.044	0.055	0.014	0.019
NO-Adv	6516	0.018	0.048	0.028	0.064	0.049	0.366	0.332	0.097	0.079
DE-Tha	26 304	0.097	0.000	0.016	0.113	0.116	0.412	0.424	0.180	0.174
SD-Dem	24 621	0.032	-0.008	0.015	0.053	0.074	0.061	0.062	0.006	0.002
ZA-Kru	24 818	0.044	0.043	0.036	0.110	0.102	0.179	0.189	0.045	0.042
CN-Cng	24 117	0.037	0.023	0.008	0.062	0.050	0.516	0.522	0.233	0.213
MY-PSO	26 301	0.220	0.068	0.034	0.455	0.423	0.079	0.079	0.017	0.016
AU-ASM	26 234	0.032	0.002	0.001	0.051	0.050	0.182	0.185	0.064	0.083

Table A9. Tair metrics for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

Tair (°C)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	8736	-10.124	1.210	0.528	2.991	2.729	0.975	0.977	0.562	0.608
CA-Man	6687	0.741	-2.198	-2.086	2.729	2.453	0.984	0.987	0.827	0.830
US-ARM	7821	15.445	1.671	1.538	2.402	2.048	0.968	0.977	0.864	0.897
BR-Sa3	8162	25.866	1.278	1.149	2.167	2.039	0.701	0.687	0.715	0.689
AR-Vir	6071	21.759	0.207	0.058	1.619	1.542	0.950	0.959	0.848	0.869
NO-Adv	4016	-2.264	-1.148	-1.663	2.222	2.099	0.941	0.969	0.381	0.562
DE-Tha	8766	9.333	-0.042	-0.030	1.793	1.708	0.965	0.969	0.760	0.814
SD-Dem	8208	27.119	0.888	1.096	1.682	1.732	0.954	0.956	0.923	0.894
ZA-Kru	8274	21.722	-0.710	-0.634	2.266	2.578	0.909	0.874	0.832	0.704
RU-SkP	7373	-2.722	-3.572	-4.295	4.192	4.620	0.982	0.982	0.705	0.706
CN-Cng	8035	6.780	-0.027	0.184	1.977	1.573	0.986	0.991	0.778	0.687
MY-PSO	8761	25.062	1.372	1.121	2.055	1.385	0.711	0.881	0.662	0.819
AU-ASM	8743	22.763	-0.468	-0.459	1.534	1.686	0.971	0.969	0.905	0.901

Table A10. PSurf metrics for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

PSurf (hPa)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	5779	1012.4	-2.9	-2.8	3.0	2.8	0.976	0.986	0.465	0.283
US-ARM	8065	976.3	0.6	0.5	1.3	0.6	0.977	0.996	0.576	0.929
BR-Sa3	7451	986.1	15.3	15.0	15.3	15.0	0.576	0.879	0.421	0.981
AR-Vir	6022	999.1	4.0	4.1	4.0	4.1	0.966	0.992	0.265	0.720
NO-Adv	4015	1011.3	-39.6	-39.4	39.7	39.4	0.983	0.994	0.922	0.983
DE-Tha	8766	972.3	-8.8	-8.9	8.8	8.9	0.983	0.997	0.269	0.955
SD-Dem	2210	949.6	-8.4	-8.2	8.4	8.2	0.744	0.974	0.064	0.466
RU-SkP	7375	987.5	-2.4	-2.4	2.4	2.4	0.993	0.996	0.813	0.886
CN-Cng	8035	992.0	5.2	5.0	5.9	5.6	0.867	0.865	0.284	0.203
AU-ASM	8732	945.0	-2.1	-2.2	2.9	2.9	0.892	0.929	0.130	0.590

Table A11. Qair metrics for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

Qair (kg kg ⁻¹)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	8764	0.00242	-0.00004	-0.00031	0.00040	0.00046	0.954	0.946	0.334	0.298
CA-Man	8741	0.00410	-0.00077	-0.00063	0.00106	0.00095	0.914	0.924	0.073	0.327
US-ARM	8763	0.00772	-0.00038	0.00048	0.00143	0.00105	0.903	0.949	0.184	0.364
BR-Sa3	8758	0.01538	0.00509	0.00370	0.00514	0.00375	0.113	0.215	0.018	0.070
AR-Vir	8765	0.01124	-0.00007	0.00069	0.00166	0.00140	0.811	0.862	0.126	0.297
NO-Adv	8742	0.00271	-0.00001	-0.00035	0.00038	0.00043	0.946	0.964	0.186	0.414
DE-Tha	8767	0.00573	0.00048	0.00077	0.00077	0.00086	0.941	0.969	0.107	0.444
SD-Dem	8767	0.00800	0.00077	0.00089	0.00139	0.00134	0.942	0.959	0.247	0.387
ZA-Kru	8767	0.01072	0.00106	0.00077	0.00189	0.00143	0.852	0.892	0.078	0.132
RU-SkP	7186	0.00359	-0.00008	0.00002	0.00061	0.00058	0.960	0.966	0.200	0.258
CN-Cng	8766	0.00486	-0.00041	0.00023	0.00071	0.00053	0.974	0.987	0.366	0.497
MY-PSO	8766	0.01633	0.00216	0.00205	0.00257	0.00206	0.320	0.506	0.159	0.215
AU-ASM	8765	0.00606	-0.00005	0.00084	0.00094	0.00102	0.921	0.956	0.090	0.437

Table A12. Wind metrics for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

Wind (ms^{-1})	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	6352	3.583	1.021	1.234	1.538	1.459	0.718	0.838	0.234	0.465
CA-Man	5683	3.404	-1.092	-0.251	1.270	0.687	0.644	0.792	0.184	0.504
US-ARM	7077	4.630	0.031	-0.306	1.351	1.091	0.709	0.820	0.232	0.582
BR-Sa3	7791	2.242	-0.829	-0.553	0.998	0.837	0.273	0.323	0.133	0.259
AR-Vir	5994	2.326	1.471	0.803	1.612	0.985	0.535	0.678	0.080	0.451
NO-Adv	2232	5.404	-1.217	-2.699	1.828	2.748	0.597	0.719	0.287	0.521
DE-Tha	8659	3.037	0.332	-0.155	1.090	0.901	0.613	0.671	0.074	0.336
SD-Dem	8206	2.870	1.403	0.892	1.839	1.264	0.339	0.581	0.007	0.440
ZA-Kru	8274	3.234	-1.580	-1.166	1.676	1.311	0.446	0.643	0.190	0.451
RU-SkP	5599	2.749	-0.732	-0.041	0.946	0.617	0.642	0.785	0.176	0.398
CN-Cng	8025	3.669	0.415	0.078	1.175	0.905	0.736	0.820	0.321	0.583
MY-PSO	8682	1.790	0.368	-0.321	0.951	0.763	0.129	0.328	-0.054	0.160
AU-ASM	8736	2.591	1.935	1.269	2.110	1.408	0.439	0.676	-0.080	0.501

Table A13. SWdown metrics for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

SWdown (W m^{-2})	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	8560	107.382	-19.644	-16.391	32.138	29.791	0.936	0.939	0.893	0.896
CA-Man	7458	132.364	2.851	0.400	41.355	38.663	0.926	0.930	0.883	0.891
US-ARM	7874	188.303	5.887	2.023	49.891	42.040	0.941	0.958	0.925	0.940
BR-Sa3	7691	188.557	-9.459	8.683	54.996	54.881	0.916	0.920	0.902	0.896
AR-Vir	4676	104.609	1.621	4.166	36.737	34.126	0.925	0.936	0.874	0.887
NO-Adv	3920	86.934	7.180	4.652	37.221	25.975	0.862	0.916	0.710	0.804
DE-Tha	8659	124.362	7.860	3.860	63.740	34.543	0.857	0.945	0.759	0.901
SD-Dem	7961	256.791	5.962	12.491	52.392	47.142	0.964	0.970	0.955	0.962
ZA-Kru	6726	198.571	18.201	1.109	59.145	47.306	0.929	0.946	0.912	0.927
RU-SkP	6983	130.778	-4.122	-0.976	26.017	41.382	0.961	0.930	0.934	0.871
CN-Cng	8380	164.232	6.550	8.133	36.527	37.803	0.953	0.958	0.934	0.936
MY-PSO	8606	189.032	0.059	-4.660	62.236	51.765	0.899	0.933	0.881	0.906
AU-ASM	8710	254.611	2.086	-2.195	85.343	81.440	0.919	0.927	0.887	0.893

Table A14. LWdown metrics for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

LWdown (W m^{-2})	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-ARM	8013	335.283	-7.250	-1.556	15.759	12.448	0.955	0.967	0.632	0.703
BR-Sa3	6103	417.552	16.713	2.293	19.635	9.403	0.738	0.762	0.707	0.727
NO-Adv	3581	287.692	-29.268	-28.467	33.945	31.531	0.872	0.881	0.258	0.366
DE-Tha	8765	315.184	-10.346	-2.396	17.538	13.257	0.897	0.916	0.442	0.568
RU-SkP	7280	260.923	-19.107	-20.013	23.150	23.341	0.971	0.973	0.527	0.491
CN-Cng	7189	288.124	-14.132	-11.582	17.637	14.097	0.977	0.983	0.558	0.663
MY-PSO	8728	417.274	-11.149	-1.690	18.903	9.322	0.684	0.768	0.621	0.681
AU-ASM	8727	346.601	-10.148	-6.204	16.340	10.773	0.946	0.970	0.591	0.729

Table A15. Precipitation (Rainf + Snowf) metrics, corrected using CRU totals, for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

<i>P</i> CRU (mm 3 h ⁻¹)	No. points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	8700	0.031	0.027	0.043	0.081	0.095	0.075	0.082	0.029	0.051
CA-Man	4454	0.122	0.126	0.092	0.280	0.236	0.211	0.282	0.068	0.209
US-ARM	8016	0.166	0.078	0.137	0.299	0.312	0.271	0.368	0.141	0.253
Br-Sa3	8758	0.482	0.139	0.235	0.929	1.033	0.083	0.063	0.029	0.030
NO-Adv	2132	0.055	0.072	0.158	0.141	0.208	0.206	0.277	0.069	0.188
DE-Tha	8767	0.291	0.046	-0.039	0.353	0.265	0.467	0.552	0.220	0.387
SD-Dem	8203	0.097	0.039	0.025	0.218	0.196	0.042	0.092	0.026	0.043
ZA-Kru	8266	0.133	0.119	0.100	0.317	0.283	0.223	0.263	0.105	0.106
CN-Cng	8036	0.111	0.032	0.052	0.168	0.157	0.483	0.575	0.307	0.444
MY-PSO	8764	0.661	0.103	0.088	1.239	1.172	0.078	0.131	-0.012	0.055
AU-ASM	8739	0.095	-0.001	0.016	0.149	0.149	0.182	0.246	0.052	0.121

Table A16. Precipitation (Rainf + Snowf) metrics, corrected using GPCP totals, for each FLUXNET2015 site: WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps.

<i>P</i> GPCP (mm 3 h ⁻¹)	No. of points	Mean Fluxnet 2015	MBE		MAE		Correlation		Adj. correlation	
			WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
US-Atq	8721	0.031	0.044	0.018	0.097	0.071	0.058	0.127	0.029	0.070
CA-Man	4454	0.121	0.101	0.131	0.264	0.259	0.200	0.312	0.070	0.225
US-ARM	8016	0.166	0.122	0.082	0.337	0.268	0.255	0.424	0.138	0.288
BR-Sa3	8758	0.482	0.247	0.135	1.028	0.936	0.071	0.088	0.026	0.048
NO-Adv	2132	0.055	0.131	0.085	0.195	0.142	0.148	0.385	0.051	0.268
DE-Tha	8767	0.291	-0.052	0.047	0.316	0.295	0.444	0.569	0.227	0.395
SD-Dem	8203	0.097	0.052	0.046	0.231	0.214	0.043	0.095	0.019	0.042
ZA-Kru	8266	0.133	0.097	0.107	0.302	0.287	0.192	0.292	0.085	0.099
CN-Cng	8038	0.111	0.063	0.023	0.201	0.133	0.348	0.654	0.214	0.508
MY-PSO	8764	0.661	0.087	0.101	1.229	1.180	0.072	0.139	-0.011	0.062
AU-ASM	8739	0.095	0.023	0.002	0.166	0.135	0.183	0.288	0.066	0.164

Table A17. Average metrics across all 13 FLUXNET2015 sites $\pm 95\%$ confidence intervals of the means for ERA5 versus FN2015 and WFDE5 versus FN2015 at hourly time steps (see Tables A1–A8).

Variable	No. of sites	Ave. MBE		Ave. MAE		Ave. correlation		Ave. adj. correlation	
		ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5	ERA5	WFDE5
Tair ($^{\circ}\text{C}$)	13	−0.560 ± 0.820	−0.209 ± 1.015	1.918 ± 0.472	2.190 ± 0.512	0.942 ± 0.045	0.935 ± 0.050	0.538 ± 0.077	0.528 ± 0.080
PSurf (hPa)	10	−4.0 ± 10.2	−3.9 ± 10.2	9.0 ± 8.2	9.0 ± 8.2	0.962 ± 0.036	0.961 ± 0.036	0.497 ± 0.240	0.453 ± 0.236
Qair (kg kg^{-1})	13	0.00044 ± 0.00055	0.00071 ± 0.00068	0.00099 ± 0.00042	0.00163 ± 0.00053	0.866 ± 0.136	0.853 ± 0.140	0.181 ± 0.051	0.160 ± 0.049
Wind (m s^{-1})	13	−0.088 ± 0.646	−0.088 ± 0.646	1.155 ± 0.333	1.155 ± 0.333	0.664 ± 0.107	0.664 ± 0.107	0.230 ± 0.057	0.230 ± 0.057
SWdown (W m^{-2})	13	2.019 ± 7.013	1.644 ± 4.287	48.042 ± 9.415	48.138 ± 9.554	0.926 ± 0.012	0.925 ± 0.013	0.671 ± 0.064	0.671 ± 0.064
LWdown (W m^{-2})	8	−9.897 ± 6.900	−8.715 ± 8.900	16.069 ± 5.706	16.473 ± 6.398	0.883 ± 0.099	0.887 ± 0.089	0.265 ± 0.068	0.280 ± 0.071
<i>P</i> CRU (mm h^{-1})	11	0.032 ± 0.048	0.027 ± 0.017	0.137 ± 0.094	0.134 ± 0.088	0.227 ± 0.103	0.195 ± 0.095	0.081 ± 0.048	0.069 ± 0.043
<i>P</i> GPCC (mm h^{-1})	11	0.032 ± 0.048	0.023 ± 0.010	0.137 ± 0.094	0.129 ± 0.046	0.227 ± 0.103	0.225 ± 0.105	0.081 ± 0.048	0.077 ± 0.046

No. of sites stands for the number of sites with measurements for each variable. Ave. stands for average. Adj. stands for adjusted.

Table A18. Average metrics across all 13 FLUXNET2015 sites $\pm 95\%$ confidence intervals of the means for WFDEI versus FN2015 and WFDE5 versus FN2015 at 3-hourly time steps (see Tables A9–A16)

Variable	No. of sites	Ave. MBE		Ave. MAE		Ave. correlation		Ave. adj. correlation	
		WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5	WFDEI	WFDE5
Tair ($^{\circ}\text{C}$)	13	0.310 ± 0.701	−0.269 ± 0.982	2.279 ± 0.432	2.169 ± 0.512	0.923 ± 0.060	0.937 ± 0.051	0.762 ± 0.098	0.768 ± 0.069
PSurf (hPa)	10	−3.9 ± 10.3	−3.9 ± 10.2	9.2 ± 8.2	9.0 ± 8.2	0.896 ± 0.098	0.961 ± 0.037	0.421 ± 0.201	0.700 ± 0.213
Qair (kg kg^{-1})	13	0.00060 ± 0.00094	0.00070 ± 0.00069	0.00146 ± 0.00077	0.00121 ± 0.00053	0.812 ± 0.164	0.853 ± 0.139	0.167 ± 0.060	0.318 ± 0.073
Wind (m s^{-1})	13	0.117 ± 0.689	−0.094 ± 0.646	1.414 ± 0.231	1.152 ± 0.334	0.525 ± 0.114	0.667 ± 0.103	0.137 ± 0.076	0.435 ± 0.074
SWdown (W m^{-2})	13	1.926 ± 5.564	1.637 ± 4.338	48.98 ± 9.808	43.604 ± 8.530	0.922 ± 0.020	0.939 ± 0.010	0.881 ± 0.042	0.901 ± 0.023
LWdown (W m^{-2})	8	−10.586 ± 10.907	−8.702 ± 8.905	20.363 ± 4.974	15.522 ± 6.572	0.880 ± 0.093	0.903 ± 0.077	0.542 ± 0.116	0.616 ± 0.109
<i>P</i> CRU (mm 3 h^{-1})	11	0.071 ± 0.031	0.082 ± 0.051	0.379 ± 0.245	0.373 ± 0.247	0.210 ± 0.101	0.266 ± 0.119	0.094 ± 0.064	0.172 ± 0.095
<i>P</i> GPCC (mm 3 h^{-1})	11	0.083 ± 0.050	0.071 ± 0.031	0.397 ± 0.249	0.356 ± 0.241	0.182 ± 0.086	0.307 ± 0.128	0.073 ± 0.063	0.197 ± 0.104

No. of sites stands for the number of sites with measurements for each variable. Ave. stands for average. Adj. stands for adjusted.

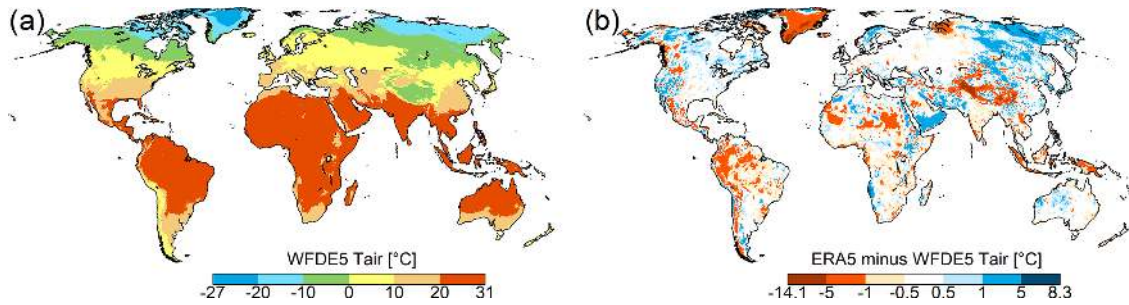


Figure A1. Long-term (1979–2016) average temperature of the climate forcings, displayed as absolute number for WFDE5 (a) and differences to ERA5, computed as ERA5 minus WFDE5 (b) (all values are given in $^{\circ}\text{C}$).

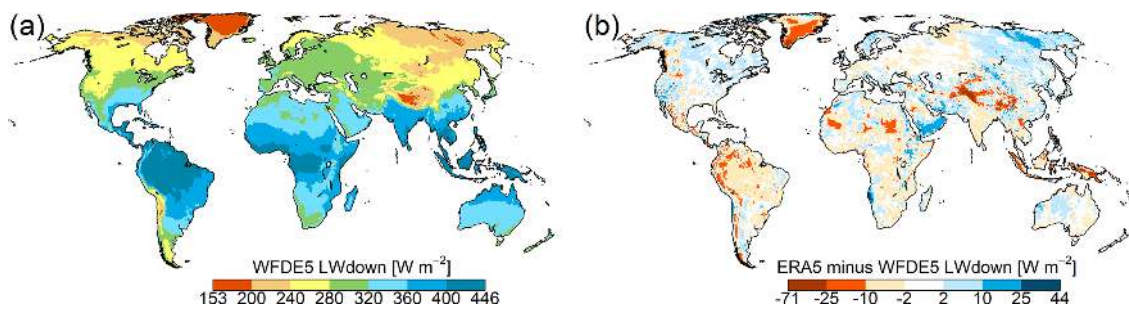


Figure A2. Long-term (1979–2016) average longwave downward radiation of the climate forcings, displayed as absolute number for WFDE5 (a) and differences to ERA5, computed as ERA5 minus WFDE5 (b) (all values are given in W m^{-2}).

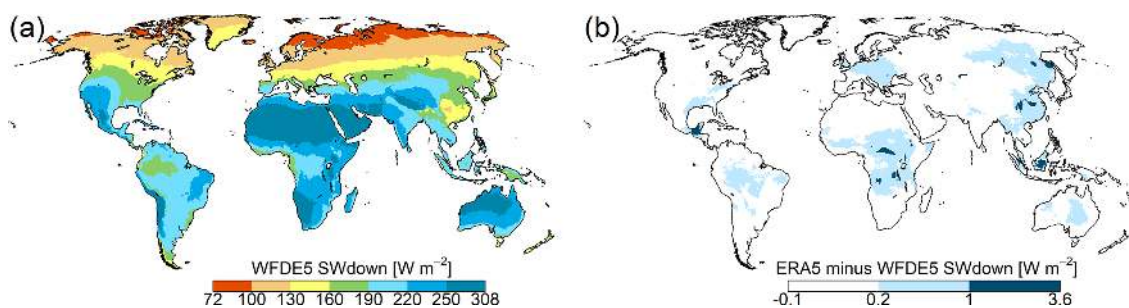


Figure A3. Long-term (1979–2016) average shortwave downward radiation of the climate forcings, displayed as absolute number for WFDE5 (a) and differences to ERA5, computed as ERA5 minus WFDE5 (b) (all values are given in W m^{-2}).

Author contributions. MC implemented the code for the generation of the dataset, led its production, and coordinated the paper's writing. GPW advised on the WATCH Forcing Data methodology and on the conversion of his WFDEI FORTRAN code into Python, ran the validations against FLUXNET2015 site observations, and checked the whole paper for consistency and English. AA contributed to the code implementation and to the production of the dataset. NB calculated the aerosol corrections. SL initiated the conversation about bias-adjusting ERA5 for impact studies; HMS highlighted the need of bias-adjusting ERA5 for hydrological applications at the ISIMIP workshop in Paris (June 2019) and validated the WFDE5 dataset with the global hydrological model WaterGAP. SL and HMS beta-tested the WFDE5 dataset. HH provided the description of the ERA5 dataset and was involved in the discussions on the creation of the WFDE5 dataset. CB had the idea of developing WFDE5, put together the team, and coordinated the different contributions. All the authors participated in the writing of the present paper, each for their own area of expertise and competence.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The WFDE5 dataset has been generated using Copernicus Climate Change Service information (2020). The contributions of ECMWF and B-Open staff toward this paper were supported by the Copernicus Climate Change Service, which is implemented by ECMWF on behalf of the European Union.

Marco Cucchi received partial support from the Department of Mathematics and Statistics of the University of Reading. All the authors thank Stefan Hagemann, Anthony Schrapffer and Xudong Zhou for their thorough and useful reviews of the manuscript. This work used surface meteorological data collected in association with the eddy covariance data acquired and shared by the FLUXNET community, including the following networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC.

Financial support. Stefan Lange received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 641816 (CRESCENDO). Hannes Müller Schmied is partly supported by the German Federal Ministry of Education and Research (BMBF, grant no. 01LS1711F).

Review statement. This paper was edited by David Carlson and reviewed by Stefan Hagemann, Anthony Schrapffer, and Xudong Zhou.

References

- Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., de Rosnay, P., and Balsamo, G.: ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better?, *Hydrol. Earth Syst. Sci.*, 22, 3515–3532, <https://doi.org/10.5194/hess-22-3515-2018>, 2018.
- Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F.: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, *Hydrol. Earth Syst. Sci.*, 21, 6201–6217, <https://doi.org/10.5194/hess-21-6201-2017>, 2017.
- Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J., Adler, R. F., and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, *Hydrol. Earth Syst. Sci.*, 23, 207–224, <https://doi.org/10.5194/hess-23-207-2019>, 2019a.
- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment, *B. Am. Meteorol. Soc.*, 100, 473–500, <https://doi.org/10.1175/BAMS-D-17-0138.1>, 2019b.
- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, *Earth Syst. Sci. Data*, 5, 71–99, <https://doi.org/10.5194/essd-5-71-2013>, 2013.
- Bellouin, N., Rae, J., Jones, A., Johnson, C., Haywood, J., and Boucher, O.: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate, *J. Geophys. Res.-Atmospheres*, 116, D20206, <https://doi.org/10.1029/2011JD016074>, 2011.
- Buck, A. L.: New Equations for Computing Vapor Pressure and Enhancement Factor, *J. Appl. Meteorol.*, 20, 1527–1532, [https://doi.org/10.1175/1520-0450\(1981\)020<1527:NEFCVP>2.0.CO;2](https://doi.org/10.1175/1520-0450(1981)020<1527:NEFCVP>2.0.CO;2), 1981.
- C3S: The Climate Data Store, available at: <https://climate.copernicus.eu/climate-data-store>, last access: 26 August 2020a.
- C3S: Near surface meteorological variables from 1979 to 2018 derived from bias-corrected reanalysis, CDS, <https://doi.org/10.24381/cds.20d54e34>, 2020b.
- Chu, H.: FLUXNET2015 dataset release, available at: <http://fluxnet.fluxdata.org/2015/12/31/fluxnet2015-dataset-release/> (last access: 26 August 2020), 2015.
- Copernicus Climate Change Service: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS), available at: <https://cds.climate.copernicus.eu/cdsapp#!/home> (last access: 26 August 2020), 2017.
- Cucchi, M., P. Weedon, G., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: Near surface meteorological variables from 1979 to 2018 derived from bias-corrected reanalysis – 2016 sample, <https://doi.org/10.21957/935P-CJ60>, last access: 26 August 2020.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, I., Biblot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Greer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de

- Rosnay, P., Tavolato, C., Thepaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. Roy. Meteorol. Soc.*, 137, 553–597, <https://doi.org/10.1002/qj.828>, 2011.
- Döll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, *J. Hydrol.*, 270, 105–134, [https://doi.org/10.1016/S0022-1694\(02\)00283-4](https://doi.org/10.1016/S0022-1694(02)00283-4), 2003.
- Ebisuzaki, W.: A Method to Estimate the Statistical Significance of a Correlation When the Data Are Serially Correlated, *J. Climate*, 10, 2147–2153, [https://doi.org/10.1175/1520-0442\(1997\)010<2147:AMTETS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<2147:AMTETS>2.0.CO;2), 1997.
- Fallah, A., Rakhshandehroo, G. R., Berg, P., O, S., and Orth, R.: Evaluation of precipitation datasets against local observations in southwestern Iran, *Int. J. Climatol.*, 40, 4102–4116, <https://doi.org/10.1002/joc.6445>, 2020.
- Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., Suzuki, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., van Vliet, M., Biber, M. F., Betts, R. A., Bodirsky, B. L., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., and Yamagata, Y.: Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), *Geosci. Model Dev.*, 10, 4321–4345, <https://doi.org/10.5194/gmd-10-4321-2017>, 2017.
- Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results, *J. Hydrometeorol.*, 12, 869–884, <https://doi.org/10.1175/2011JHM1324.1>, 2011.
- Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, *J. Hydrometeorol.*, 12, 556–578, <https://doi.org/10.1175/2011JHM1336.1>, 2011.
- Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, *Sci. Data*, 7, 109, <https://doi.org/10.1038/s41597-020-0453-3>, 2020.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, *Earth Syst. Dynam.*, 4, 219–236, <https://doi.org/10.5194/esd-4-219-2013>, 2013.
- Hersbach, H., Peubey, C., Simmons, A., Berrisford, P., Poli, P., and Dee, D.: ERA-20CM: A twentieth-century atmospheric model ensemble, *Q. J. Roy. Meteorol. Soc.*, 141, 2350–2375, 2015.
- Hersbach, H., Bell, W., Berrisford, P., Horányi, A., J., M.-S., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., and Dee, D.: Global reanalysis: goodbye ERA-Interim, hello ERA5, <https://doi.org/10.21957/VF291HEHD7>, 2019.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Lalouaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, *Q. J. Roy. Meteorol. Soc.*, online first, <https://doi.org/10.1002/qj.3803>, 2020.
- Hirahara, S., Alonso-Balmaseda, M., de Boisseson, E., and Hersbach, H.: Sea Surface Temperature and Sea Ice Concentration for ERA5, available at: <https://www.ecmwf.int/node/16555> (last access: 26 August 2020), 2016.
- Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, *Atmos. Chem. Phys.*, 19, 3515–3556, <https://doi.org/10.5194/acp-19-3515-2019>, 2019.
- Jones, P. W.: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates, *Mon. Weather Rev.*, 127, 2204–2210, [https://doi.org/10.1175/1520-0493\(1999\)127<2204:FASOCR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2), 1999.
- Lange, S.: Bias correction of surface downwelling longwave and shortwave radiation for the EWEMBI dataset, *Earth Syst. Dynam.*, 9, 627–645, <https://doi.org/10.5194/esd-9-627-2018>, 2018.
- Lange, S.: Earth2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), GFZ Data Services, <https://doi.org/10.5880/pik.2019.004>, 2019a.
- Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), *Geosci. Model Dev.*, 12, 3055–3070, <https://doi.org/10.5194/gmd-12-3055-2019>, 2019b.
- Lange, S.: WFDE5 over land merged with ERA5 over the ocean (W5E5), GFZ Data Service, <https://doi.org/10.5880/pik.2019.023>, 2019c.
- Lange, S.: ISIMIP3BASD v2.4.1, Zenodo, <https://doi.org/10.5281/zenodo.3898426>, 2020.
- Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M., and Döll, P.: Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, *Hydrol. Earth Syst. Sci.*, 18, 3511–3538, <https://doi.org/10.5194/hess-18-3511-2014>, 2014.
- Müller Schmied, H., Adam, L., Eisner, S., Fink, G., Flörke, M., Kim, H., Oki, T., Portmann, F. T., Reinecke, R., Riedel, C., Song, Q., Zhang, J., and Döll, P.: Variations of global and continental water balance components as impacted by climate forcing uncertainty and human water use, *Hydrol. Earth Syst. Sci.*, 20, 2877–2898, <https://doi.org/10.5194/hess-20-2877-2016>, 2016.
- Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6), 1970.
- New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century Space–Time Climate Variability. Part I: Development of a 1961–90 Mean Monthly Terrestrial Climatol-

- ogy, *J. Climate*, 12, 829–856, [https://doi.org/10.1175/1520-0442\(1999\)012<0829:RTCSTC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0829:RTCSTC>2.0.CO;2), 1999.
- New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century Space–Time Climate Variability. Part II: Development of 1901–96 Monthly Grids of Terrestrial Surface Climate, *J. Climate*, 13, 2217–2238, [https://doi.org/10.1175/1520-0442\(2000\)013<2217:RTCSTC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2), 2000.
- Pastorello, G. Z., Papale, D., Chu, H., Trotta, C., Agarwal, D. A., Canfora, E., Baldocchi, D. D., and Torn, M. S.: A new data set to keep a sharper eye on land–air exchanges, *Eos*, 98, 28–32, <https://doi.org/10.1029/2017EO071597>, 2017.
- Sato, M., Hansen, J. E., McCormick, M. P., and Pollack, J. B.: Stratospheric aerosol optical depths, 1850–1990, *J. Geophys. Res.-Atmos.*, 98, 22987–22994, <https://doi.org/10.1029/93JD02553>, 1993.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: GPCP’s new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theor. Appl. Climatol.*, 115, 15–40, <https://doi.org/10.1007/s00704-013-0860-x>, 2014.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., and Ziese, M.: GPCP Full Data Monthly Product Version 2018 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data, Global Precipitation Climatology Centre, https://doi.org/10.5676/DWD_GPCP/FD_M_V2018_050, 2018.
- Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling, *J. Climate*, 19, 3088–3111, <https://doi.org/10.1175/JCLI3790.1>, 2006.
- Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T.: A global space-based stratospheric aerosol climatology: 1979–2016, *Earth Syst. Sci. Data*, 10, 469–492, <https://doi.org/10.5194/essd-10-469-2018>, 2018.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Anderson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, I., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, *Q. J. Roy. Meteor. Soc.*, 131, 2961–3012, <https://doi.org/10.1256/qj.04.176>, 2005.
- Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project framework, *P. Natl. Acad. Sci. USA*, 111, 3228–3232, <https://doi.org/10.1073/pnas.1312330110>, 2014.
- Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: The WATCH Forcing Data 1958–2001: A meteorological forcing dataset for land surface- and hydrological-models, Tech. rep., WATCH Technical Report 22, available at: <http://www.eu-watch.org/publications/technical-reports> (last access: 26 August 2020), 2010.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, *J. Hydrometeorol.*, 12, 823–848, <https://doi.org/10.1175/2011JHM1369.1>, 2011.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resour. Res.*, 50, 7505–7514, <https://doi.org/10.1002/2014WR015638>, 2014.
- Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate Res.*, 30, 79–82, <https://doi.org/10.3354/cr030079>, 2005.
- scheischler, J., Fischer, E. M., and Lange, S.: The effect of univariate bias adjustment on multivariate hazard estimates, *Earth Syst. Dynam.*, 10, 31–43, <https://doi.org/10.5194/esd-10-31-2019>, 2019.