Response to reviewer #1:

Dear Graham,
We appreciate the time taken on reviewing our manuscript, and the comments you have made. Our response to each comment are given in bold font below:

Minor comments:
1) Although the authors target updating beyond the current coverage of GPCCv7 precipitation (2013) using GPCC products, in fact datasets such as WFDEI already extend to the end of 2015 through the use of the CRU observations of precipitation. Therefore the need for updating is not as severe as implied, especially for Europe where the half-degree resolution GPCC and CRU precipitation totals show good overall matches. Additionally, the post 2013 data (CRU of Rainf_WFDEI_CRU plus Snowf_WFDEI_CRU) could be used within the validation of the new product. The reason why CRU precipitation for 2014 and 2015 has not been utilized should be made clear.

It is true that the CRU-version of WFDEI is updated more frequently than appears from our manuscript, and we will make statements to emphasise the more freuqent updates. We have not used any data after 2013 because we wanted to have a complete set of data for all data sets, and we believe the WFDEI product based on GPCC data is the better one globally as well as more compatible with the current GFDCL version. Extending the analysis beyond 2013 would therefore introduce more even more complexity in the paper with multiple versions of WFDEI. We will therefore not extend the analysis beyond 2013.

2) The abstract should make it clear that the new product is available at half-degree resolution and for daily precipitation and near-surface air temperature only (other variables provided by the existing datasets, such as downwards shortwave radiation fluxes [suitable for land surface models and global hydrological models] are not provided).

We now mention that on several occasions in the revised version.

3) The name GFD (Global Forcing Data) is very generic. This name does not convey the fact that only temperature and precipitation are involved nor that the data are updated to near real time. I would suggest a change to something like Current Global Forcing Data to emphasise the value added.

Point well taken, and we have on several occasions considered a different name of the product. It has however been used in many applications already and it would be confusing for current users to re-name it now. We have after some thought made a name change to the overall method to HydroGFD, because hydrological applications is the main aim of this data set. This also separates it more clearly from the WFD data set which could otherwise be confusing.

4) The updating methodology is dependent on the availability of ERA Interim products. In the next couple of years ERA Interim will no longer be available and a different reanalysis (ERA-5) will be provided by ECMWF instead. The nature of ERA-5 is such (higher resolution, multiple realizations) that a smooth transition from ERA-Interim to ERA-5 for the GFD is not guaranteed. Some comment on this would be appropriate.

GFD is generic in the sense that it can easily be applied to other re-analysis and observational data sets. We will comment on our plans for switching to ERA-5 once enough data has been released in 2018, and that we will then rebuild the complete data product to make use of the

full ERA-5 data.

Figure changes:
Fig 2 is currently difficult to comprehend. The caption says "Climatological mean (top) precipitation" - the brackets should say "(left)" similarly "and (bottom) temperature" - the brackets should say "(right)". The top row shows absolute means and the colour bars should be provided next to them. The next three rows show differences and the other colour bars should be there. However, the caption says: "the relative difference EI, GFDCL, and WFDEI." This should be changed to "(left) the relative difference in precipitation and (right) absolute difference in temperature." Finally to be clear what is shown every panel should have its own heading. For example, panel b should be headed with something like: "100%x(EI minus GPCCv7)/GPCCv7".

We appologize for the incorrect information in this figure caption due to a last minute restructuring. We have made changes both to the caption and the figure itself to make everything more clear.

Fig 6 It is noticeable that the authors have been careful to avoid poor colour pallets that could confuse colour-blind readers. However, it is very hard to distinguish the yellow shades as well as the blue shades for the E-HYPE maps. Can the colour scheme be changed to show clearer gradations?

We have adapted the colorbrewer palettes and have thereby restricted ourselves to few and clearly distinguishable colors for all plots.

Fig 7 The labelling is misleading. It is far easier for the reader to understand this figure if every panel has a heading of either "Europe" or "Arctic". Also every Y axis should have the evaluation metric indicated for every panel ("Bias", "NSE", "r", "Variability"). The caption should also spell out NSE, r (is this Pearson’s or Spearman’s?) and what is meant by "variability" (is this standard deviation or variance?).

We have clarified the figure and restructured it along with the other hydrological evaluation figures to follow the same structure.

Minor text corrections:
p4 line 14 The word "data" is plural. A dataset (singular) contains a lot of data (plural). Hence in both places on this line change "is" > "are".
p4 line 14 "On notable" > "One notable".
p5 line 33 and p6 line 1 "Priestly" > "Priestley".
p8 line 5 "method overestimate" > "method overestimates".
p8 line 5 "The updating method also produce" > "The updating method also produces".
p8 line 7 "a difference already in the observations" > "a difference in the observations".
p8 line 15-16 "is mainly used interim to bridge" > "is mainly used as an interim measure to bridge".
p10 line 6 "for the first about 90-100" > "for about the first 90-100".
p10 line 11 "which is on a much larger magnitude" > "which is of a much greater magnitude" [note the "of" not "on"].

Thank you for the detailed language checks, which are much appreciated! We will correct accordingly.

Response to anonymous reviewer #2:

We appreciate very much the reviewer's comments which we will answer below.

Major comments:
1.) In the introduction the authors claim that "forcing data for large scale hydrological models is essentially not available. . ." – This claim is not correct. In fact there are numerous, (semi) operational global hydrological models that have solved the problem of forcing data in different ways, for example the Global Flood Awareness System, the Global Flood Monitoring System, and the Global Flood Forecasting System. Please refer to their relevant scientific articles on what type of global forcing data they use to derive hydrologic model initial conditions. In addition, see also the recent article from Hirpa et al. (AMetSoc, 2016) on this topic which should be considered by the authors. Finally, also ERA-5 is now available and produced in "near real time". This should also be inclued and discussed.

Such claims are of course only subjectively correct, and we will rewrite that statement. From our point of view, we need consistent data for temperature and precipitation that are close to observations. This issue has been solved in different ways for different projects with a global approach. Each of the listed systems above have made their own versions of forcing data, but they do not share the data themselves openly as far as we are aware, thus not "available". We will discuss these different global systems and how they have solved the issue in the revised introduction.
We will also discuss gridded and satellite products that can be used as forcing, including their pros and cons.
Thanks for the Hirpa et al (2016) paper, which we have traced to the paper Hirpa, F. A., Salamon, P., Alfieri, L., Pozo, J. T. D., Zsoter, E., & Pappenberger, F. (2016). The effect of reference climatology on global flood forecasting. *Journal of Hydrometeorology*, *17*(4), 1131-1145. We will include the results in the revised introduction.
Regarding ERA-5, we intend to make use of this once available for a long time period during 2018, and make a new version of GFD based on that. We add this to discussions and outlook.

2.) Numerous datasets that are claimed to lack temporal coverage have in fact coverage also of recent years including the MSWEP dataset. The manuscript should reflect the latest status of these datasets.

Our intention with temporal coverage was for the "near real-time", which to us is until at least last month. We will mention the full extent of WFDEI-CRU and MSWEP.

3.) Furthermore, the authors have omitted completely the TRMM (and now Global Precipitation Measurement) datasets that represent an important source of near-real time precipitation forcing which is clearly the most important variable for the forcing data of hydrological models. Those need to be at least mentioned/referenced with an explanation of why those are not used in this work.

We have added a section in the introduction where we discuss the different data sets with daily resolution data and a global or near global extend, including CPC-products, TRMM and GPM.

4.) The authors claim that their method is similar to the method used in the WFDEI dataset. Yet, in Fig. 2d the relative difference between WFDEI and GPCC is considerable whereas the relative difference between GFDCL and GPCC is very small. This suggest that the changes introduced by the authors in comparison to the methodology to WFDEI have a significant impact. Instead of claiming that this is simply due to the use of different precipitation sources this should be further investigated and explained.

The method is very similar, but the version of GPCC is different between WFDEI and GFDCL, which makes all the difference. We have more clearly distinguished the methodology and the data set itself in the revised version. The methodology is the same as in WFDEI, but the data sets are different, which explains the differences in the end product.

5.) This manuscript has almost 20 (!) abbreviations. Some of them are spelled out before their first use, others not. Some important ones such as GFDOD1 and GFDOD2 are not properly explained. Even though I am familiar with most of the abbreviations it makes this manuscript very hard to read. The authors should at least include a table with an overview of the most important ones (maybe expand tables 1 and 2) or maybe add them as an annex.

We have tried various versions of this, and found the current to be most clear in combination with brevity in the text. We will consider other ways to describe the data, and to perhaps remove some abbreviations. The GFDOD1 and GFDOD2 refers to the two separate months of GFDOD, however, as there are only minor differences between these months, we will remove them from the paper completely and only discuss GFDOD.

6.) What is the effect on hydrological simulations when you update with the new observational data on the 10th of each month? This might lead to a significant discrepancy between simulations done on the 9th and then, with the updated dataset on the 10th. Clearly that represents an issue for hydrological forecasting but is not properly discussed by the authors.

This is indeed an issue. We present this issue in more detail in the revised version, and describe that the last 3-month history of a simulations changes after the update and that forecasts are not "compatible" in the overlapping window of the forecasts around that date.

7.) Figure 6 d seems to suggest that there is actually less or at least similar bias in the average upstream runoff difference when compared to 6b and 6c. This seems to contradict Figure 5 where the OD period shows the highest absolute difference. Please explain more in detail why this is the case.

We have changed the palette so that the bias is more clearly visible, which better emphasizes the differences here.

8.) The GFD claims to be a global dataset for hydrological models. Yet, the hydrological validation was only performed for catchments in northern latitudes. There is currently no hydrological validation for basins located in tropical climates. The validation should be improved including also basins from these regions.

We have only made detailed simulations for the two norther regions, however, there are operational forecasts made with HYPE initialised with GFD for more tropical regions, such as Niger. We will link to such operational systems and their performance when published.

9.) The manuscript lacks a paragraph on future developments.

A short outlook for 2018 was added, with addition of new data sets as well as moving to ERA-5 instead of ERA-Interim.

10.) The title of the manuscript should be modified and the authors should define in the text what they mean with “near real time”. “Near real time” suggests that data is updated within hours or at least days and not monthly.

Near real-time is of course in general not well defined and depends on the application. No system can every be “real-time” altough some describe their systems that way. The word “near” therefore works as an alert to the reader that it might not be up to their standards of what “real-time” is. We will therefore keep the title, as it points out the added value of GFD in comparison to WFDEI which is more epidodically updated. To mention but one similar definition; in the precentation of GLOFAS in Alfieri et al. (2013), ERA-Interim is introduced as near real-time data, with its three month delay.

Below, we silently accept all changes unless commented.

Minor comments:

- Please add the datasets used for WFDEI to Table 2

- Please add a reference for GHCN-CAMS into the references

Its already there: Fan and van den Dool (2008)

- P.10, line 14: last sentence is unclear. Please describe further and rephrase

- What is the difference between GFDOD1 and GFDOD2?

We remove these as described above.

- The nonlinear scale in Fig. 2 and 4 makes it very difficult to look at the results. Basically everything greater than +-25% is hardly distinguishable. Please choose a different scale or use the one applied in Fig. 6 (and possible also a different color scale)

We have revised all the figures for better color palettes, now using colorbrewer suggestions, and moved to a linear scale for most plots.

- In the section on “Meteorological evaluation” the authors write “. . .and focus instead on comparisons to the WFDEI dataset.” However, in the following evaluation you compare the GFDCL mostly with GPCC, EI or OD. Please clarify.

- Does Fig. 2 show the relative difference of EI, GFDCL and WFDEI to GPCC7/CRUts? Please make this more clear.

- Is the precip bias between EI and GPCC7 in line with other studies looking at the precip bias of EI? If yes please add the relevant reference.

We do not see what such a reference would add. Our method is clearly defined and there is no room for other studies showing different results.

Near real-time adjusted reanalysis forcing data for hydrology

Peter Berg[1], Chantal Donnelly[1], and David Gustafsson[1]

[1]Hydrology Research Unit, Swedish Meteorological and Hydrological Institute, Folkborgsvägen 17, 601 76 Norrköping, Sweden

Correspondence to: Peter Berg (peter.berg@smhi.se)

Abstract. [c1]Extending climatological forcing data to current and real-time forcing is a necessary task for hydrological forecasting. While such data are often readily available nationally, it is harder to find fit-for-purpose global data sets that span long climatological periods through to near real-time. Hydrological simulations are generally sensitive to bias in the meteorological forcing data, especially relative to the data used for the calibration of the model. The lack of [c2]high quality daily resolution data at a global scale has previously been solved by adjusting re-analysis data [c3]with global gridded observations. However, existing data sets of this type have been produced for a fixed past time period determined by the main global observational data sets. Long delays between updates of these data sets leaves a data gap between [c4]the present [c5]day and the end of the data set. Further, hydrological forecasts require initialisations of the current state of the snow, soil, lake (and sometimes river) storage. This is normally conceived by forcing the model with observed meteorological conditions for an extended spin-up period, typically at a daily time step, to calculate the initial state. Here, we present [c6]and evaluate a method named [c7]HydroGFD (Hydrological Global Forcing Data to combine different data sets in order to produce near real-time updated hydrological forcing data [c8]of temperature and precipitation that are compatible with the products covering the climatological period. [c9]HydroGFD resembles the already established WFDEI method (Weedon et al., 2014) closely, but uses updated climatological observations, and for the near real-time it uses interim products that apply similar methods. This allows [c10]HydroGFD to produce updated forcing data including the previous calendar month around the 10^{th} of each month. We present the [c11]HydroGFD method and [c12]therewith produced data sets, which are evaluated against [c13]global data set[c14]s, as well as with hydrological simulations with

[c1] ~~Updating climatological forcing data to near current data are compelling for impact modelling, e.g. to update model simulations or to simulate recent extreme events.~~

[c2] *Text added.*

[c3] *Text added.*

[c4] *Text added.*

[c5] *Text added.*

[c6] *Text added.*

[c7] ~~GFD (Global Forcing Data)~~

[c8] *Text added.*

[c9] *Text added.*

[c10] *Text added.*

[c11] *Text added.*

[c12] ~~different~~

[c13] ~~the WFDEI~~

[c14] *Text added.*

the HYPE model over Europe and the Arctic region[c15]s. We show that [c16]HydroGFD performs similarly to WFDEI and that the updated period significantly reduces the bias of the reanalysis data[c17]. For real-time updates until the current day, extending [c18]HydroGFD with operational meteorological forecasts, a large drift is present in the hydrological simulations due to the bias of the meteorological forecasting model.

1 Introduction

[c1]Large scale hydrological models at global or continental scales require meteorological forcing data at, typically, daily time resolution. There is a lack of data with high quality and consistency between variables at such scales, however, data at coarser monthly scales are more prominent. Reanalysis data fulfill the spatial and temporal consistency, but suffer from bias that limits their use for hydrological simulations. Current data sets that merge reanalysis and coarser observations [c2]bridge the data gap, but are mostly only episodically updated (Sheffield et al., 2006; Weedon et al., 2011, 2014; Beck et al., 2016).

The degree to which the skill of a hydrological forecast is sensitive to the initial hydrological conditions on one hand, and the meteorological forcing in the forecast period on the other hand, depends on factors such as the hydro-meteorological regime of the catchment and the memory of the hydrological system. The hydrological skill sensitivity to the initial state and/or the meteorological forecast varies as a function of the season, which have been shown for both seasonal and short term forecasts (Li et al., 2009; Shukla and Lettenmaier, 2011; Paiva et al., 2012; Demirel et al., 2013; Pechlivanidis et al., 2014). In most cases, however, hydrological forecast models are initialized by hindcast simulations covering some period before the forecast issue date, for which appropriate meteorological forcing data are needed.

Climatological hydrological simulations require consistent forcing data for a long period[c3], which can be problematic with gauge based data sets if the gauge location and the network density are very different between the observed variables. [c4]Observational data sets with global coverage are sparse regarding data with at least daily resolution, but there are exceptions such as the Climate Prediction Center's (CPC) products for temperature (CPCtemp, 2017)[c5] and precipitation (Chen et al., 2008)[c6]. There are also several promising satellite based products, such as the TRMM (Huffman et al., 2009b)[c7] and GPM missions, although satellite data require adjusments to ground truth observations. The negative aspects of the above data sets are problems

[c15] *Text added.*

[c16] *Text added.*

[c17] ~~, although less well for the last two months of the updating cycle~~

[c18] *Text added.*

[c1] ~~Forcing data for large scale hydrological models are essentially not available in near real-time, and gridded observational data sets with near global coverage are generally only available at too coarse time steps.~~

[c2] ~~are not available with regular updates, and often lack coverage of the recent years~~

[c3] *Text added.*

[c4] *Text added.*

[c5] *Text added.*

[c6] *Text added.*

[c7] *Text added.*

with spatial coverage, by non-sampled (polar) regions for the satellite data and lack of gauges in parts of the world for gauge based data as the gauge density becomes even more important at the daily time scale.

[c8]

Operational models working on a global scale have found ways to work with sparse observations. The Global Flood Awareness System (GloFAS) uses the ERA-Interim reanalysis (Dee et al., 2011), with precipitation adjusted using data from GPCP (Huffman et al., 2009a) at a monthly time-scale (Alfieri et al., 2013; Hirpa et al., 2016). Another global scale model system is the Global Flood Forecasting Information System (GLOFFIS), where the meteorological forcing data is derived from several sources, such as gauge measurements, CPC-Unified gridded precipitation (Chen et al., 2008) and the ECMWF control forecast (Emerton et al., 2016).

Earlier methods (Sheffield et al., 2006; Weedon et al., 2011, 2014; Beck et al., 2016) have merged information from a re-analysis with temporally coarser observational data, to produce new data sets that inherit the temporal resolution of the re-analysis with the average properties of the observations. With these methods, long periods of [c1]internally consistent daily or sub-daily resolution and global coverage become available for, e.g., large scale hydrological simulations. The various methods have applied different re-analysis data sets and observational records, and therefore differ in their final result. The more simple method is that of Weedon et al. (2014), where mainly single data sets are applied globally for the adjustment of each variable. Although this leaves the method highly dependent on the quality and availability of few data sets, it makes the method less affected by temporal and spatial inconsistencies between periods and regions. [c2]An issue with relying on gridded observational data sets is that such data are often updated episodically, and with several months or even years of delay before they are updated. This can be an issue for global or continental hydrological forecasting where up-to-date information is important, thus requiring a continuous updating of the forcing data while retaining a consistent climatology.

Here, we present the [c3]HydroGFD [c4]method for producing [c5]adjusted meteorological forcing data sets for a near global domain.The novelty in the production of the data sets is the combination of reanalysis and operational global model input, as well as the combination of various observational data sources to fill the gap between the present and the end of the climatological products. We evaluate the updating procedure to the climatological data by direct comparison of the meteorological data, as well as by employing a hydrological model to evaluate the data sets. The main motivation for creating the data set is to update climatological simulations, but also to improve the initialisation for hydrological forecasting at large scales or in data sparse regions where dense observational data are not available for initialisation. We present evaluation of two such applications for the Arctic and European set-ups of the hydrological models E-HYPE and Arctic-HYPE.

[c8] ~~Global observational data are often updated in discrete steps, and with several years delay. This can be an issue for impact modelling where up-to-date information is important, thus requiring a continuous updating of the forcing data, while retaining a consistent climatology.~~

[c1] *Text added.*

[c2] *Text added.*

[c3] *Text added.*

[c4] ~~(Global Forcing Data)~~

[c5] ~~scaled~~

Table 1. Table of meteorological forcing data used in the analyses and hydrological simulations. The data sets are described in Tab.2.

Abbreviation	Atm. model	Precipitation	Wet days	Temperature	Period
GFDCL	EI	GPCC7	CRU	CRU	1979–2013
GFDEI	EI	GPCC-Monitor	GPCC-FG daily	GHCN-CAMS	2010–(t-3 months)
GFDOD	OD	GPCC-FG monthly	GPCC-FG daily	GHCN-CAMS	2010–(t-1 month)
OD	OD	NA	NA	NA	2010–t

2 Methods and Data

The [c6]HydroGFD method is currently intended to be a substitute and extension of [c7]precipitation and temperature from the WFDEI method (Weedon et al., 2011), which is currently used in many hydrological simulations with HYPE (Lindström et al., 2010) and other hydrological models.

We are therefore mimicking the WFDEI set-up closely, however, with some necessary differences due to updates of the meteorological observations since the first appearance of WFDEI. [c1]The HydroGFD data set is currently limited to precipitation and temperature at three and six hourly intervals, wheras WFDEI produces several additional variables (Weedon et al., 2011). The basic method is to[c2] construct monthly mean adjustment factors [c3]per calendar month for each variable and to adjust every time step during the month with that factor. For temperature, the adjustment factor is produced by subtracting the monthly mean reanalysis from the observations, and adding this to every time step of the reanalysis. For precipitation, a first step of adjusting the number of wet days is performed. The underlying assumption is that the reanalysis model produces excessive light rainfall (drizzle). Days with the least amount of rainfall that are in excess to the observed rainy days are set to zero. In a second step, the ratio between the monthly mean observations and the reanalysis data is calculated and used to scale the reanalysis data.

The [c4]HydroGFD system has been applied to produce the main climatological dataset called GFDCL, which is [c5]a methodological equivalent to the WFDEI (Weedon et al., 2011) dataset except for updated climatological observations (see Tab. 1) and differences in the implementation. GFD[c6]CL, like WFDEI, is based on the ERA-Interim (EI) reanalysis but is coded so that EI can be interchanged with other reanalyses. Precipitation is corrected for wet day bias compared to CRUts3.22 wet day information, and scaled with monthly precipitation from GPCC7 (see Tab. 2). Temperature was corrected additively with CRUts3.22 monthly mean temperature. The GFD[c7]CL data set is restricted to the time period 1979–2013, due to the start of the EI reanalysis period, and by the end of the GPCC7 (Schneider et al., 2014) observational data set. The main difference

[c6] *Text added.*
[c7] *Text added.*
[c1] ~~GFD is adjusting precipitation and temperature at three and six hourly fields, respectively.~~
[c2] ~~, for each calendar month,~~
[c3] *Text added.*
[c4] *Text added.*
[c5] ~~an~~
[c6] *Text added.*
[c7] *Text added.*

Table 2. Table of model and data sources used in the analyses.

Data set	Variables	Resolution	Period	Reference
ERA-Interim (EI)	T, P	~0.8°	1979–(t-3 months)	Dee et al. (2011)
ECMWF-OD (OD)	T, P	~0.22°	2010–t	
CRUts3.22 (CRU)	T, P, wet-day*	0.5°	1901–2013	Harris and Jones (2014)
GPCC7	P	0.5°	1901–2013	Schneider et al. (2015b)
GPCC-Monitor(v5)	P	1.0°	1982–(t-2 months)	Schneider et al. (2015a)
GPCC-FG	P**, wet-day***	1.0°	2009–(t-1 month	Ziese et al. (2011); Schamm et al. (2013)
GHCN-CAMS	T	0.5°	1948–(t-1 month)	Fan and Van den Dool (2008)
WFDEI	P****, T*****	0.5degree	1979-2013(6)	Weedon et al. (2011)

* Gridded from SYNOP stations. ** Using the GPCC First guess monthly product

*** Derived from daily time-step information from the GPCC first guess daily product.

**** Using different versions of GPCC until 2013, also a version using CRU until 2016.

***** Using different versions of CRU until 2016.

[c8]between GFDCL and WFDEI arises from the treatment of under-catch, i.e. the rainfall likely not captured by the rain gauges due to turbulence around the gauge. WFDEI applied the Adam and Lettenmaier (2003) under-catch correction to the GPCC5 and GPCC6 data sets. With GPCC7, under-catch correction is already included in the data set, and need not be applied in the [c9]HydroGFD methodology. However, for GPCC7, the under-catch correction was based on Legates and Willmott (1990), but reduced by 15% to better fit with their own estimates (Schneider et al., 2014). Adam and Lettenmaier (2003) compared their method with that of Legates and Willmott (1990) and found the latter to lead to too low precipitation amount by about 5–30%, and differences in the annual cycle of the correction factors. There is clearly a large controversy [c10]on this topic. We therefore expect differences between GFDCL and WFDEI in both annual totals and in the annual cycle.

[c1]The main issue tackled here is how to implement the WFDEI methodology forward in time as GPCC7 becomes unavailable, or when EI becomes unavailable. We propose two flavours of [c2]HydroGFD [c3]to extend the period past year 2013 (see Tab. 1 for data sets and references):

1. GFDEI consists of the EI data set with precipitation scaled by the GPCC monitoring data set and wet day adjusted according to the GPCC first guess daily product. Temperature is adjusted with the GHCN-CAMS data set.

[c8] ~~of~~

[c9] *Text added.*

[c10] ~~in how much under-catch to expect~~

[c1] ~~T~~

[c2] *Text added.*

[c3] ~~are produced~~

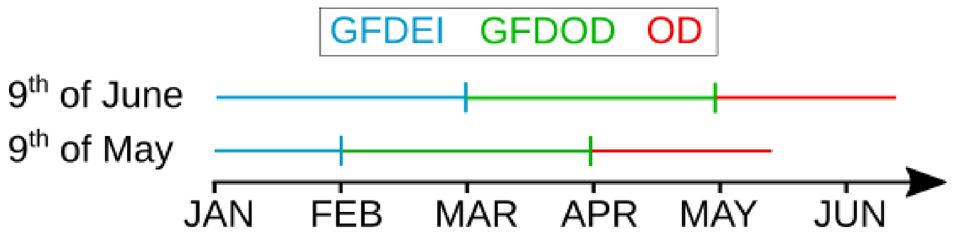

Figure 1. Schematic of the updating procedure. The [c3]Hydro GFD data are continuously updated with GFDEI as long as EI data are available. The intermediary data set GFDOD fills up the time series as long as GPCC data are available, and then continues with uncorrected OD data. Because the previous month becomes updated on the 10^{th} of each month, the 9^{th} is the day with the longest period of OD driving data. The next month, GFDEI is extended one month, and the GFDOD data are updated for the new month.

2. GFDOD consists of the ECMWF deterministic forecast, which differs from EI by mainly the model version and the assimilated data. Precipitation is scaled by the GPCC first guess monthly data set and wet day adjustments according to the GPCC first guess daily product. Temperature is adjusted with GHCN-CAMS data.

GFDEI fills the gap between the end of GFDCL in 2013 until the latest available EI data, i.e. until about three months ago. For the last two months, GFDOD is used to fill the gap. The necessary datasets are all available for download around the 10^{th} in each month. Fig. 1 shows a schematic [c1]for how the forcing data is used to update hydrological models to today's date. [c2]For example, to update a model to the 9^{th} of May, the model is forced with GFDEI until the 31^{st} of January, GFDOD until 31^{st} of March then OD until the 9^{th} of May. This gives a period of 40 days with unadjusted OD data. However, to update the model to the 10^{th} of May, because the GPCC monitoring product becomes available on the 10^{th} of the month (at latest) all data shifts one calendar month and require a shorter period of OD data (unadjusted data). In a hydrological forecasting context, the simulations are updated from the GFDEI data, which is the continuous extension of GFDCL, and the GFDOD and OD parts are re-run after each update to determine the new initial conditions.

Because the observational data sets only provide information over land areas, the [c4]HydroGFD system only produces adjustments where data [c5]are available, and retains the original reanalysis, or deterministic forecast, when no data is available. On[c6]e notable exception is Antarctica, which is not covered by the observational data sets, and is therefore not adjusted at any step of the updating procedure.

[c1] ~~for the update procedure~~

[c2] ~~E.g., on the 9^{th} of May, the available GFDEI data extends until the end of January and the GFDOD data until the end of March. Thereafter, the ECMWF deterministic model, without adjustments, fills the gap until the present day. Therefore, the 9^{th} is the day of the month with the longest period of non-adjusted data. At the 10^{th} of May, new observational data arrive, and both GFDEI and GFDOD are extended by one month. The period of non-adjusted data then shrinks to ten days.~~

[c4] *Text added.*

[c5] ~~is~~

[c6] *Text added.*

2.1 HYPE model

The HYPE (Hydrological Predictions for the Environment) model is a process[c7] based hydrological model developed for high-resolution multi-basin applications, which has been applied at various spatial scales (from tens to million square kilometres) and hydro-climatological conditions (Lindström et al., 2010; Strömqvist et al., 2012; Arheimer et al., 2012; Andersson et al., 2015; Gelfan et al., 2017). The model is based on a semi-distributed approach where the hydrological system is represented by a network of sub-basins, which are further divided into classes that can be selected to represent combinations of soil-type and land-cover or elevation zones. The water balance and runoff from each sub-class is calculated taking into account processes such as snow and glacier accumulation and melt, infiltration, evapotranspiration, surface runoff, tile drainage and groundwater recharge and runoff. The runoff from the land classes is further routed through the network of lakes and rivers represented by the sub-basin delineation. The model is used for research and operational purposes, to provide information for, for instance, flood and hydro power reservoir inflow forecasting, river discharge and nutrient loads to the ocean, as well as assessment of climate change impact on hydrological systems.

To evaluate the real usefulness of the [c1]HydroGFD data in [c2]continental (and by extension global) hydrological forecasting, the [c3]HydroGFD data was tested in two continental scale applications of HYPE. For Europe, the E-HYPE v3.2 (Hundecha et al., 2016) hydrological model was calibrated with GFDCL and employed to evaluate the updating versions of [c4]HydroGFD. The simulation domain ranges from wet Arctic, wet maritime to dry Mediterranean climatic conditions. The E-HYPE model has been shown to reproduce well the spatial and temporal variability in hydrological processes across Europe (Donnelly et al., 2016; Hundecha et al., 2016), and has been identified as a useful model for continental scale forecasting (Emerton et al., 2016). E-HYPE takes daily mean precipitation and temperature as input. Potential evapotranspiration is estimated from daily mean temperature and extraterrestrial radiation estimated separately for each sub-basin location and day of the year using the modified Jensen-Haise/McGuiness model following Oudin et al. (2005). For each sub-basin, air temperature and precipitation is taken from the nearest [c5]grid point. Temperature is further corrected with a constant lapse rate (-0.65 $^{\circ}C/100\ m$) for the difference between the mean sub-basin elevation and the corresponding elevation of the [c6]grid point. Elevation correction of precipitation is also possible in the HYPE model, but it is not used in E-HYPE.

For the Arctic, we use the Arctic-HYPE model v3.0 (Andersson et al., 2015; Gelfan et al., 2017) that covers the land area draining into the Arctic Ocean (excluding Greenland). The model domain is 23 million km^2 divided into 32599 sub-basins with an average size of 715 km^2. The Arctic region is characterized by numerous lakes of various size (5% areal fraction) and glaciers (about 50% of the glaciated area outside the Greenland and Antarctica Ice sheets, mainly on islands in the Canadian Arctic archipelago, Svalbard, and Russian Arctic islands) (Dyurgerov and Meier, 1997; Meier and Bahr, 1996). To take into

---

[c7] ~~ed~~

[c1] *Text added.*

[c2] *Text added.*

[c3] *Text added.*

[c4] *Text added.*

[c5] ~~GFD~~

[c6] ~~GFD~~

account the long turnover times of larger lakes in the domain (for instance Lake Baikal) and the on-going decline in glacier volume, the Arctic-HYPE model was initialized using an initial spin-up period for the period 1961–2010 using the WFD data (Weedon et al., 2011) with a simplified correction of precipitation versus GPCC7 on a monthly basis, to be consistent with the GFDCL data, and extended using GFDCL for the period 1979-2013. As for E-HYPE, Arctic-HYPE is forced by daily mean precipitation and temperature, but in contrast to E-HYPE, potential evapotranspiration is calculated using the Priestl[c7]ey-Taylor equation assuming it to be more representative for the wide range of climatic conditions in the Arctic-HYPE domain. The Priestl[c8]ey-Taylor equation requires solar radiation and relative humidity, which was estimated using the minimum and maximum daily temperatures as additional input variables, following the recommended procedures by Allen et al. (1998).

Both E-HYPE and Arctic-HYPE models have been parametrized and calibrated with similar step-wise approaches involving first of all sub-basin delineation based on globally available digital elevation data (USGS HydroSHEDS and Hydro1K). Secondly, classification into selected land-use and soil type classes based on land cover and soil data such as the ESA CCI Land cover or CORINE, and HWSD respectively. Thirdly, model parameters governing water balance processes in ice/snow, soil, lakes and rivers were thereafter calibrated in an iterative procedure using river discharge data from the Global Runoff Data Center (GRDC), as well as data on internal water balance components such as snow (ESA GlobSnow and Former Soviet Union Snow course data), glaciers (glacier area and mass balance data from ESA CCI Glacier and the World Global Monitoring Service), and evapotranspiration (fluxtower data from FluxNet and MODIS Evaporation products).

For the evaluation simulations with [c1]HydroGFD products, the models are run once per month from 9^{th} of May 2010 to 9^{th} of December 2013, to recreate a 130 day initialization simulation for each run, ending on the given date. This is the longest possible initialization step, as the meteorological forcing data are updated at the 10^{th}, for which the initializations would advance one calendar month (Fig. 1). The first simulation starts from a saved state of the GFDCL simulation in January 2010, and each subsequent run is initialised from a starting state saved from the GFDEI portion of the previous simulation; making the GFDEI simulation continuous in time. A total of 44 simulations are made with each hydrological model. The simulations are then compared with a climatology simulated using GFDCL forcing for each region for the same period 2010–2013 [c2]to evaluate the change in simulated hydrology as a result of the changing forcing data products.

3 Results

We begin with evaluating the GFDCL data set, as well as comparing differences between the various [c3]HydroGFD versions. Thereafter, we present analysis of hydrological simulations for Europe and the Arctic.

[c7] *Text added.*
[c8] *Text added.*
[c1] *Text added.*
[c2] *Text added.*
[c3] *Text added.*

3.1 Meteorological evaluation

Climatology 1979–2013: GFDCL is directly comparable to the WFDEI data set due to the very similar method[c4], but will differ due to different underlying data, and handling of precipitation under-catch. Because WFDEI was on several occasions evaluated against flux tower measurements across the globe (Weedon et al., 2011, 2014; Beck et al., 2016), we do not repeat such evaluation for GFDCL here, and [c5]compare instead [c6] to the WFDEI [c7]and other data set[c8]s.

The baseline reanalysis data set EI has both wetter and drier regions compared to GPCC7, with biases towards ±100% [c1]over large regions (Fig. 2b). Overall, the wetter regions are [c2]predominant. Here, we note especially the wet bias throughout the Arctic (excluding Greenland), and mainly slightly wet bias in continental Europe. Corrections with GFD[c3]CL reproduces GPCC7 well (Fig. 2c), as [c4]expected per definition of the method. There are some isolated patches with underestimated precipitation, mainly in the dry regions of the Sahara desert and southern Arabic Peninsula, which appears because no scaling is possible for single months with a complete lack of precipitation in EI at these locations. In contrast to GFDCL, WFDEI has a general wet bias when compared to GPCC7 (Fig. 2d). The wet bias is explained mainly by stronger under-catch corrections included in WFDEI, as explained in Section 2.

Temperature bias in EI [c5]ranges mainly between $\pm1\ ^{\circ}C$ for most land areas (Fig. 2f), but there are regions with considerable bias. There is a mostly warm bias of partly several degrees Celsius in the Arctic regions. Europe has a low bias, except for Scandinavia, which shows a warm bias. Both GFDCL (Fig. 2g) and WFDEI (Fig. 2h) correct the bias per definition, and are both indistinguishable at [c6]the $0.2\ ^{\circ}C$ accuracy [c7]of the color legend, even though different versions of CRU were employed (GFDCL: CRUts3.22; and WFDEI:CRUts3.1 for 1979–2009, CRUts3.21 for 2010–2012, and CRUts3.23 for 2013).

In summary, GFDCL is methodologically similar to WFDEI and differences in the results are mainly due to the different precipitation source used.

Evaluation of the updating method (2010–2013): To evaluate the updating method of the GFDEI and GFDOD datasets, we investigate differences in bias for the period 2010–2013 when all data sources are available (see Tab. 2). The only methodological difference between GFDEI/OD and GFDCL is the calculation of the number of wet days in a month. Whereas the latter uses gridded station measurements of the number of wet days from CRU[c8], the former data sets have the number of wet days calculated from the GPCC-FG daily product as the number of days in a month with precipitation larger than or equal to

[c4] *Text added.*
[c5] ~~focus~~
[c6] ~~on comparisons~~
[c7] *Text added.*
[c8] *Text added.*
[c1] ~~that cover~~
[c2] ~~dominating~~
[c3] *Text added.*
[c4] ~~practically~~
[c5] ~~is low ($<0.5\ ^{\circ}C$)~~
[c6] ~~$0.25\ ^{\circ}C$~~
[c7] *Text added.*
[c8] ~~ts3.22~~

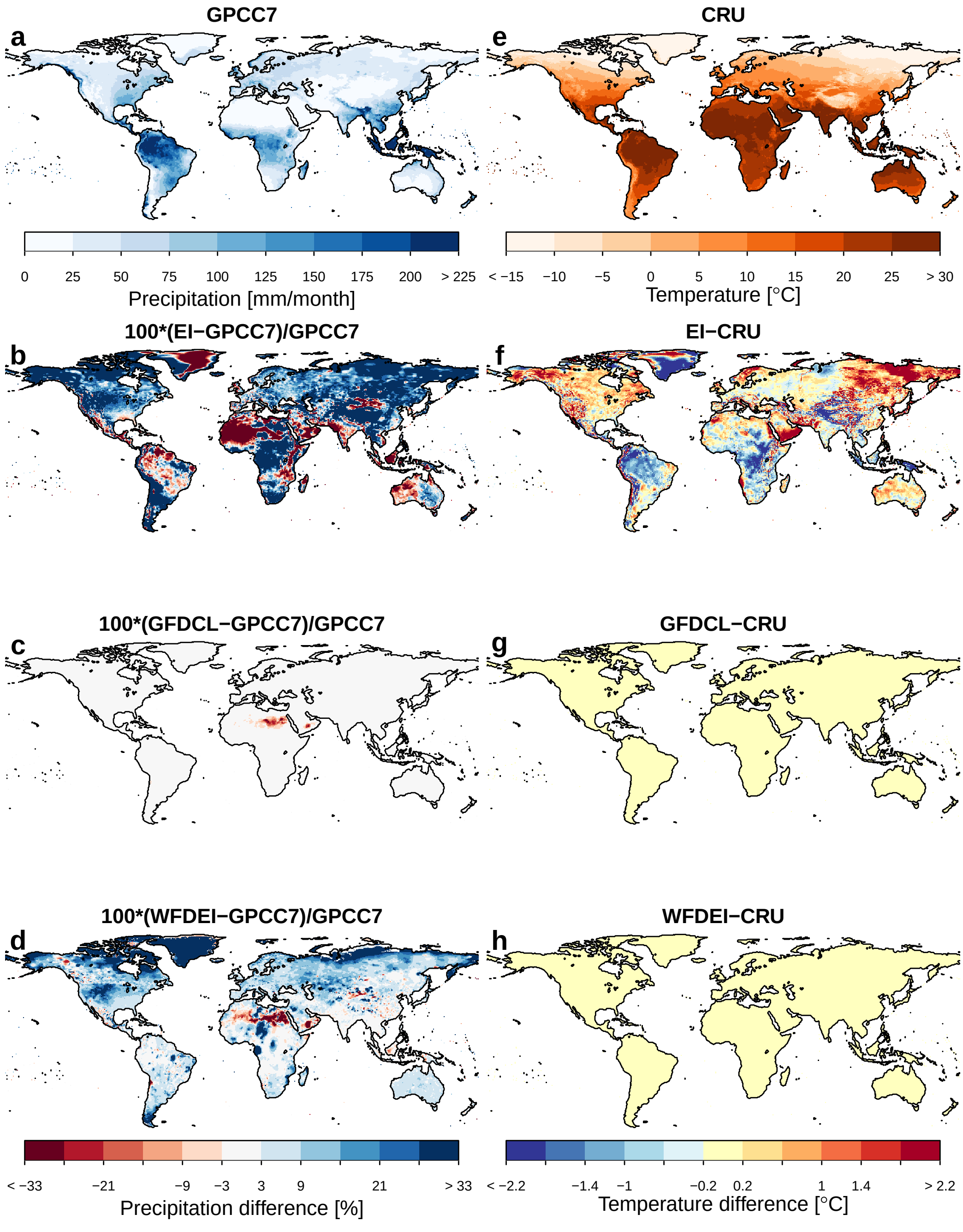

Figure 2. Climatology of (a) precipitation from GPCC7, and (e) temperature for CRU. Relative difference in climatological precipitation from GPCC7 for (b) EI, (c) GFDCL and (d) WFDEI. Absolute difference in climatological temperature from CRU for (f) EI, (g) GFDCL and (h) WFDEI.

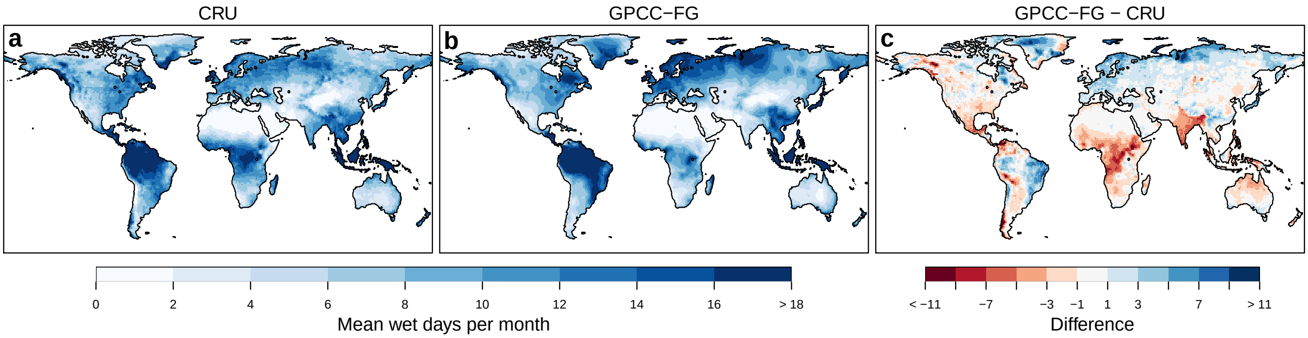

Figure 3. Comparison of the number of wet days provided by (a) the CRU[c1] data set, compared to those derived from (b) GPCC-FG, and (c) the difference between the two for the period 2010–2013.

1 mm/day. Fig. 3 presents the period average number of wet days in a month for CRU[c9] and GPCC-FG. The two methods to calculate wet days differ significantly for Europe and especially the Arctic part of Scandinavia and western Russia, where the updating method overestimate[c10]s the number of wet days. The updating method also produce[c11]s underestimations in Africa, Latin America and the Andes. An interesting difference is markedly confined within the political borders of India, which implies a difference [c12]in the observations entering either CRU[c13] or GPCC-FG, and could be an artefact of a higher station density in that region compared to surrounding regions [c14]or a different threshold used for the wet-day definiton.

Fig. 4 shows the bias between the different data sets used here, such that the data set given at the top of the plot is compared with that named to the left of each row. In the first row (Fig. 4a–d), all data sets are compared to GPCC7. Clearly, GPCC-Monitor and GPCC-FG both underestimate precipitation for most parts of the globe compared to GPCC7. This is partly due to the lack of under-catch correction, but differences may also result from lower station density, as not all stations are available in real-time. The latter effect can be seen in the different bias patterns for GPCC-Monitor and GPCC-FG (Fig. 4a and b, respectively), and also in the difference between GPCC-Monitor and GPCC-FG (Fig. 4e). The extension of the GFDCL data set is mainly through the GFDEI product, which is adjusted by GPCC-Monitor, and the GFDOD product is mainly used [c2]as an interim [c3]measure to bridge the data gap for initializations of forecasts. GFDEI has a similar spatial structure as GPCC7, with some marked regional differences, but a general reduction of a few percent in total precipitation is seen. EI has a similar bias as for the climatological period (compare Fig. 4c and Fig. 2b). The bias of GPCC-Monitor shrinks in significance when compared to that of EI, which means that the extension of GFDCL with GFDEI is indeed relevant when extending the climatological data set for, e.g., hydrological applications.

[c9] ~~ts3.22~~
[c10] *Text added.*
[c11] *Text added.*
[c12] ~~already~~
[c13] ~~ts3.22~~
[c14] *Text added.*
[c2] *Text added.*
[c3] *Text added.*

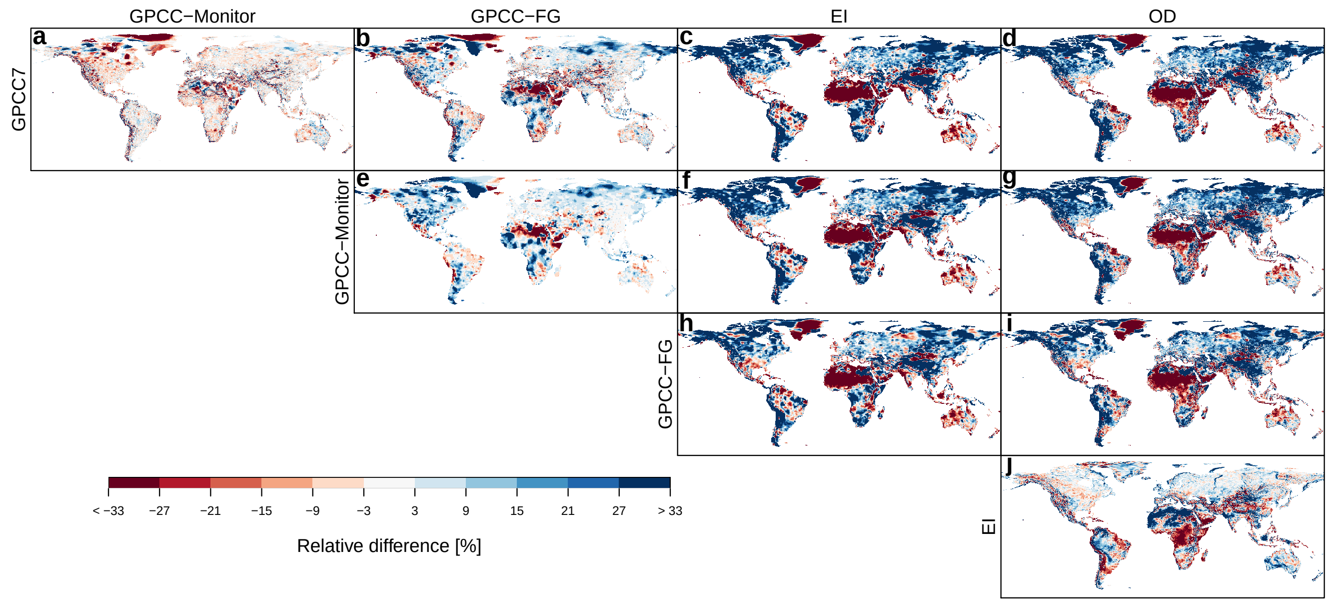

Figure 4. Relative difference of mean monthly precipitation between different data sources and (a–d) GPCC7, (e–g) GPCC-Monitor, (h–i) GPCC-FG, and (j) EI.

OD has a similar bias as EI when compared to GPCC7 (Fig. 4d), however, also clear differences although of lower magnitude appear in a direct comparison of OD and EI (Fig. 4j). The main differences are confined to the tropical regions, however, the bias of OD is much more prevalent than that of GPCC-FG, which indicates value in the interim GFDOD product. GFDEI and GFDOD retains the average bias of the GPCC-Monitor and GPCC-FG products, per definition (not shown).

Temperatures are compared between the data sets GHCN-CAMS, EI and OD toward CRU[c1] (not shown). The main differences are in the Arctic, especially for Greenland, and for various mountain ranges and coastal areas, with magnitudes of several degrees Celsius. EI and OD have similar bias for most of the globe, although OD has a larger warm bias in the Arctic and northern Europe.

3.2 Hydrological evaluation

[c2]The effect of the interim products on simulated hydrology in Europe and the Arctic are evaluated using the E-HYPE and Arctic-HYPE continental hydrological models. The resulting bias at the end of OD simulation is indicative of the potential bias in initial conditions for a hydrological forecast made using the HydroGFD procedure. First, a climatological simulation driven by GFDCL is carried out for the years 2010–2013, starting from a saved model state the $10^{t}h$ of January 2010. Second, a set of simulations separated by one calendar month was carried out for the period 10^{th} of May 2010 until 10^{th} of November

[c1] ~~ts3.22~~

[c2] ~~Evaluation simulations are performed both for E-HYPE and Arctic-HYPE in the following procedure.~~

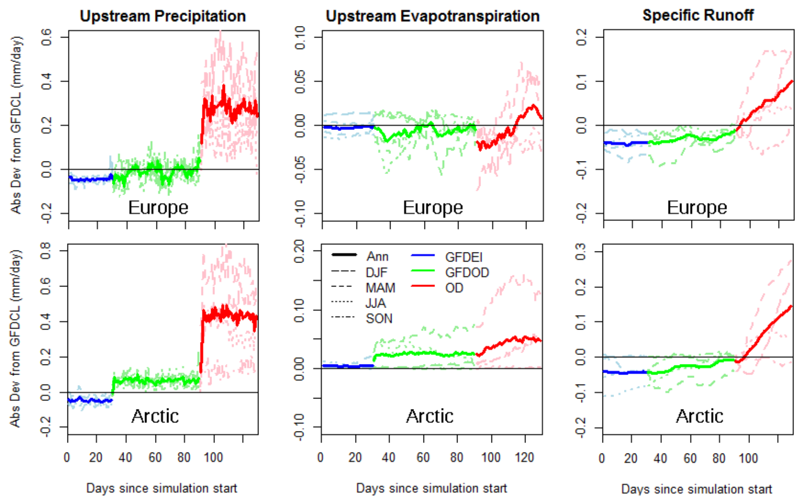

Figure 5. Upstream precipitation, evapotranspiration and specific runoff averaged over all catchments and shown for all forecast times as well as per season for (top) E-HYPE and (bottom) Arctic-HYPE. All runs are presented as deviations from the GFDCL forced simulation.

2013. Each of the simulations start from GFDEI for the first month, continue with GFDOD for two months, and then OD for one month and ten days (see Fig. 1). The model state of the last day of the GFDEI simulation is saved and used for the initial state of the next month's GFDEI simulation. When nothing else is stated, the evaluation is performed with day one at the first day of the GFDEI until the last day of the simulation, which is approximately day 130. In the figures we mark with colours as in Fig. 1 the different forcing data periods approximated by 30 day months to indicate which data set was used.

The impact of the differences in the GFDEI, GFDOD and OD data sets compared to the reference GFDCL simulation are shown as an average across the respective simulation domains in Fig. 5. The specific runoff shows lower values for GFDEI and GFDOD [c1]compared to GFDCL for both domains. Clearly, the main determining factor for the differences arise from the differences in upstream precipitation from the first 30 days with GFDEI. Even though GFDOD has less of [c2]a precipitation offset from GFDCL, and for the Arctic even a positive difference, the GFDEI offset causes a slow drift [c3]in runoff toward the new conditions of GFDOD, and therefore a remaining negative offset for [c4]about the first [c5] 90–100 days. Upstream evapotranspiration shows a low offset from GFDCL for GFDEI, which shows that the GHCN-CAMS and CRUts3.22 data sets are similar for these two domains. However, although the same data set is used for GFDOD, there is a larger offset for this period. The difference in upstream evapotranspiration offsets between the two model domains is most likely a result of the larger (and positive) offset in upstream precipitation for the GFDOD and OD periods in the Arctic-HYPE domain, rather

[c1] *Text added.*
[c2] ~~an~~
[c3] *Text added.*
[c4] *Text added.*
[c5] ~~about~~

than the smaller differences in temperature. OD has a strong wet precipitation bias [c6]<u>(particularly in the northern hemisphere; results for the tropics and southern hemisphere may be different)</u>(Fig. 4d), which is [c7]<u>of</u> a much [c8]<u>greater</u> magnitude than that of GFDEI. The bias causes the slow drift of the specific runoff to accelerate around day 90–100, as the model adjusts to the new precipitation average. The case is similar for both domains. Another striking feature from Fig. 5 is the larger variability for GFDOD and OD, compared to GFDEI, which is due to differences between EI and OD. This affects the [c9]<u>day-to-day variations</u> of the simulations, but not the [c10]<u>total</u> water balance.

Fig. 5 shows also results per season. For both Europe and the Arctic, precipitation and runoff biases are largest for the OD forced period in DJF and MAM, and relatively minor in JJA and SON. Seen as a continental mean, there is little variation in the biases between individual years, meaning that the results are robust in time (not shown).

Fig. 6 shows a spatial view [c1]<u>of</u> the average upstream runoff difference from the GFDCL simulation for each domain. [c2] In the resolution of the colour scales, there [c3]<u>are</u> only small differences between GFDEI and GFDOD[c4]. The offsets from GFDCL are mainly within ±20% for Europe, but much stronger local offsets are seen in the Arctic domain. The Arctic is a more sensitive region to differences in the station density behind the gridded observational data sets, as there are fewer stations to begin with. This fact plays a large role in shaping the offsets seen here. The OD period is, as expected, wetter for most of the domains, but more clearly so for the Arctic domain.

A selection of in-situ observations from gauging stations with available data from at least two of the four simulated years was used [c5]<u>to analyse how the model performance against observed discharge varies using the climatological forcing and different interim data sets</u>. Performance criteria of the models for each of the gauges are presented for each data set in comparison to GFDCL in Fig. 7. Since GFDCL is always the reference, the results for each gauge lines up vertically in the figure. The two domains show similar results, and we therefore describe the results in a general sense. The bias follows the patterns described above, with lower values for GFDEI and GFDOD, while OD has higher values. Whether there is a positive o[c6]<u>r</u> negative bias is determined by the initial bias of the GFDCL simulation. NSE and [c7]<u>Pearson</u> correlation (r) are not showing any clear structure, but remain reasonable for most of the simulations. The [c8]<u>variance</u> is consistently higher for the OD simulation as also noted above.

[c6] *Text added.*

[c7] ~~on~~

[c8] ~~larger~~

[c9] ~~dynamics~~

[c10] *Text added.*

[c1] *Text added.*

[c2] ~~The two months with GFDOD were here divided into two separate maps to see if there is any difference dependent on the time elapsed since the GFDEI simulation.~~

[c3] ~~is~~

[c4] ~~1, and even less between GFDOD1 and GFDOD2~~

[c5] ~~for the analysis~~

[c6] ~~f~~

[c7] *Text added.*

[c8] ~~variability~~

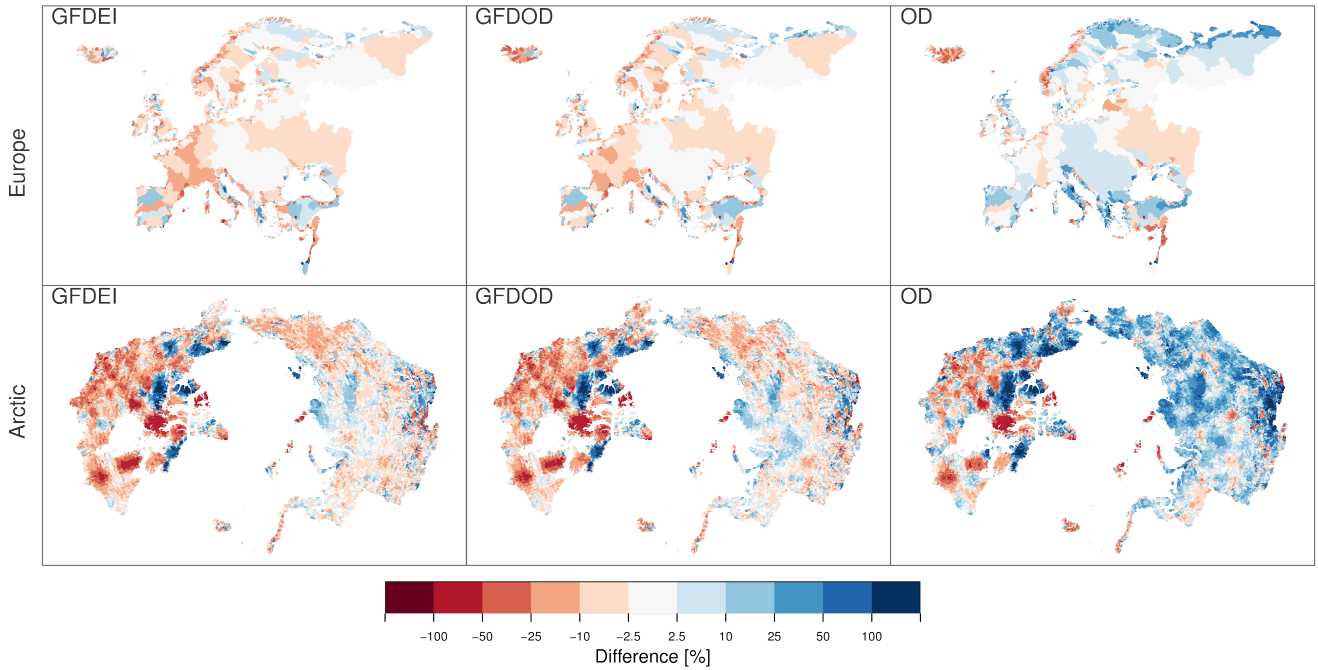

Figure 6. Relative upstream specific runoff difference from GFDCL for each catchment of (top) E-HYPE and (bottom) Arctic-HYPE, with the different data sets (right to left) GFDEI, GFDOD and OD.

In summary, the domain average deviations from GFDCL shows that the updating procedure adds value to the simulations by keeping the precipitation and temperature climate closer to the GFDCL data set [c11]when compared to the alternative of using uncorrected data (e.g. OD). The extension of GFDCL with GFDEI has [c12]only minor effects on the long term hydrology. However, for forecast initializations, the inevitable switch to OD data closer to the "current" date, i.e. the day to issue a forecast, there is a strong drift due to the wet bias of OD [c13]in the northern hemisphere regions evaluated here. The drift continues throughout the OD period, which means that the initial drift a forecast is subjected to is dependent on the day of the forecast. The drift is largest for forecasts issued just before the 10^{th} and lowest just after. [c14]This warrants future development to look for a method to adjust the deterministic forecast data (OD). In highly seasonal regions with little interannual variability, OD could be adjusted with the monthly climatological mean precipitation and temperature; however, it should be investigated whether this worsens simulations in regions with high interannual variability. Such a correction could also be used within the forecasting period; however is reserved as the subject of future study.

[c11] *Text added.*
[c12] ~~therefore~~
[c13] *Text added.*
[c14] *Text added.*

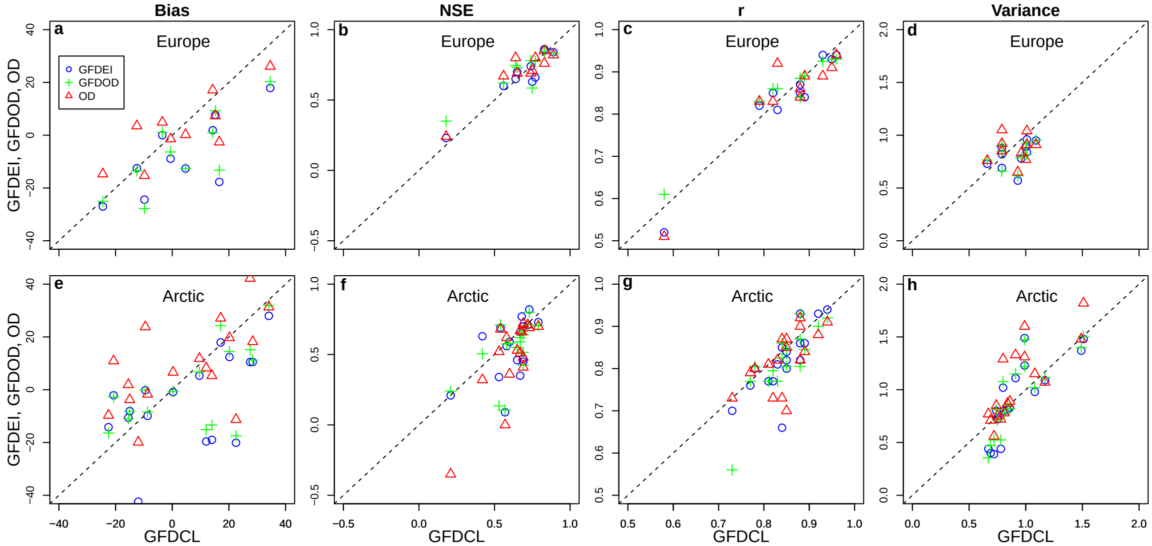

Figure 7. River discharge model performance measures: bias (relative volume error in %), Nash-Suthcliff Efficiency (NSE), Pearson correlation (r), and ratio of simulated and observed variance for a selection of grid points in [c9] (top) Europe, and (bottom) Arctic. [c10]The performance of GFDEI, GFDOD, and OD (y-axis) is compared to GFDCL (x-axis) in scatter diagrams.

4 Conclusions

We present [c1]and evaluate a new data set called HydroGFD, which consists of several [c2]interim products to fill the gap between available climatological and forecasted data. The main product, GFDCL, is the [c3]methodological equivalent to the already well established WFDEI (Weedon et al., 2014), although with updated observational data sets. To extend the data set beyond year 2013, when e.g. the GPCC7 data set ends, adjustments are performed with regularly updated data sets. This is performed with the GFDEI product until the latest update of EI, which is with about a three month delay. For near real-time updates, GFDOD makes use of the ECMWF deterministic model with similar data sets for adjustments as for GFDEI. GFDOD is available until the end of the previous month from around the 10^{th} of the current month.

GFDCL is found to be a much similar product [c1]to WFDEI, but with a more consistent data set. The introduced under-catch corrections in the precipitation data set GPCC7 differ from that assumed in WFDEI, which leads to generally lower amounts in GFDCL. Temperature is very similar.

[c1] *Text added.*
[c2] ~~products.~~
[c3] *Text added.*
[c1] ~~as~~

The updates in GFDEI beyond 2013, are evaluated for an overlapping period (2010–2013). GFDEI is found to have slightly lower precipitation amounts, and spatially somewhat different temperatures. However, the differences to GFDCL shrinks in comparison to the bias of EI which has bias of often an order of magnitude higher.

When EI is not available, the OD model is employed and the precipitation data source changes from GPCC-Monitor to GPCC-FG. The change in data source has the largest impact, with several geographical differences which impact on the GFDOD product. As an interim product until the next update, GFDOD reduces the bias of OD (which is similar to that of EI) to levels similar to GFDEI.

Initializations of hydrological simulations for forecasting purposes are investigated for GFDOD, extended by the non-corrected OD until the day before the next update of GFDOD. It is found that the strong bias of OD, especially for precipitation, causes a severe drift of the hydrological model away from the GFDOD climatology. The results are similar for both the domains investigated, i.e. Europe and the Arctic region. Some measure to reduce the induced drift due to bias of OD would be necessary for reliable forecasts. [c1]Further, as HydroGFD data are updated, it is necessary to re-run the hydrological model from the last update of EI, i.e. for the last three months. The effect of the updating procedure will be that the forecast just after the update, will not be consistent with the one from the day before due to the change in the last few months and the initial state at the time of the forecast. Analysis of the forecasts was not part of the current study.

[c2]HydroGFD is currently applied for forecasts with HYPE models in the Niger river basin (http://hypeweb.smhi.se/nigerhype/forecast/) which is evaluated in (Andersson et al., 2017)[c3], the Arctic (http://hypeweb.smhi.se/arctichype/forecast/), as well as for seasonal forecasts in a concept study for Copernicus Climate Change Services available from the sectoral information services at http://climate.copernicus.eu/.

The HydroGFD data sets are planned for public release via a web interface on http://hypeweb.smhi.se/. [c4] [c5]An updated version of HydroGFD using the new reanalysis system ERA-5 and introducing further observational data sets is foreseen during 2018.

5 Data availability

The [c6]HydroGFD method relies mainly on open data sets, as referenced within the article. ECMWF reanalysis can be accessed via the web portal https://www.ecmwf.int/en/research/climate-reanalysis/era-interim/. The forecasts from ECMWF (here referred to as "OD"), are restricted to member institutes (or other special circumstances, see https://www.ecmwf.int/en/forecasts/accessing-forecasts), and are therefore not available for public download. However, [c7]HydroGFD will shortly appear online on http://hypeweb.smhi.se/. Hydrological simulations were performed with the open source model HYPE, which can be accessed at http://hypecode.smhi.se/.

[c1] *Text added.*
[c2] *Text added.*
[c3] *Text added.*
[c4] ~~Besides the climatological data set GFDCL, the website will feature monthly updates of the GFDEI and GFDOD data sets.~~
[c5] *Text added.*
[c6] *Text added.*
[c7] ~~the GFDCL, GFDEI and GFDOD data sets, with regular updates,~~

Acknowledgements. We acknowledge the hard work of building the data sets used within the presented work. This includes the data from the ERA-Interim, CRU, GPCC and WFDEI as referenced within the paper, as well as GHCN-CAMS (National Center for Atmospheric Research Staff (Eds). Last modified 08 May 2014. "The Climate Data Guide: GHCN (Global Historical Climatology Network) Related Gridded Products." Retrieved from https://climatedataguide.ucar.edu/climate-data/ghcn-global-historical-climatology-network-related-gridded-products.) and the ECMWF deterministic forecast system. Further, we acknowledge the inital work of implementing the [c1]HydroGFD system at SMHI by Lisa Bengtsson, Magnus Lindskog and Heiner Körnich, and the work on operationalization by Fredrik Almén.

[c1] *Text added.*

References

Adam, J. C. and Lettenmaier, D. P.: Adjustment of global gridded precipitation for systematic bias, J. of Geophys. Res. Atm., 108, 2003.

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F.: GloFAS - global ensemble streamflow forecasting and flood early warning, Hydrol. Earth Syst. Sci., 17, 1161–1175, doi:10.5194/hess-17-1161-2013, 2013.

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., et al.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, FAO, Rome, 300, D05 109, 1998.

Andersson, J., Pechlivanidis, I., Gustafsson, D., Donnelly, C., and Arheimer, B.: Key factors for improving large-scale hydrological model performance, European Water, 49, 77–88, 2015.

Andersson, J. C., Ali, A., Arheimer, B., Gustafsson, D., and Minoungou, B.: Providing peak river flow statistics and forecasting in the Niger River basin, Physics and Chemistry of the Earth, Parts A/B/C, 100, 3–12, doi:10.1016/j.pce.2017.02.010, 2017.

Arheimer, B., Dahné, J., Donnelly, C., Lindström, G., and Strömqvist, J.: Water and nutrient simulations using the HYPE model for Sweden vs. the Baltic Sea basin - influence of input-data quality and scale, Hydrol. res., 43, 315–329, 2012.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25°global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data, Hydrol. Earth Syst. Sci. Discuss., pp. 1–38, doi:10.5194/hess-2016-236, 2016.

Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R., and Janowiak, J. E.: Assessing objective techniques for gauge-based analyses of global daily precipitation, J. Geophys. Res. Atm., 113, D04 110, doi:10.1029/2007JD009132, 2008.

CPCtemp: https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globaltemp.html, 2017.

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, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quart. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Demirel, M. C., Booij, M. J., and Hoekstra, A. Y.: Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models, Water Resour. Res., 49(7), 4035–4053, doi:10.1002/wrcr.20294, 2013.

Donnelly, C., Andersson, J., and Arheimer, B.: Using flow signatures and catchment similarities to evaluate a multi-basin model (E-HYPE) across Europe, Hydr. Sciences Journal, 61(2), 255–273, doi:10.1080/02626667.2015.1027710, 2016.

Dyurgerov, M. and Meier, M.: Year-to-year fluctuations of global mass balance of small glaciers and their contribution to sea-level changes, Arct. Alp. Res., 29(4), 392–402, 1997.

Emerton, R. E., Stephens, E. M., Pappenberger, F., Pagano, T. C., Weerts, A. H., Wood, A. W., Salamon, P., Brown, J. D., Hjerdt, N., Donnelly, C., and Baugh, C. A.: Continental and global scale flood forecasting systems, WIRES Water, 3(3), 391–418, 2016.

Fan, Y. and Van den Dool, H.: A global monthly land surface air temperature analysis for 1948–present, J. Geophys. Res. Atm., 113, D1, 2008.

Gelfan, A., Gustafsson, D., Motovilov, Y., Kalugin, A., Krylenko, I., and Lavrenov, A.: Climate change impact on the water regime of two great Arctic rivers: modeling and uncertainty issues, Clim. Chang., 141 (3), 1–17, doi:10.1007/s10584-016-1710-5, 2017.

Harris, I. and Jones, P.: CRU TS3.22: Climatic Research Unit (CRU) Time-Series (TS) Version 3.22 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901- Dec. 2013), doi:10.5285/18BE23F8-D252-482D-8AF9-5D6A2D40990C, 2014.

Hirpa, F. A., Salamon, P., Alfieri, L., del Pozo, J. T., Zsoter, E., and Pappenberger, F.: The Effect of Reference Climatology on Global Flood Forecasting, J. Hydrometeorol., 17, 1131–1145, doi:10.1175/jhm-d-15-0044.1, 2016.

Huffman, G. J., Adler, R. F., Bolvin, D. T., and Gu, G.: Improving the global precipitation record: GPCP Version 2.1, Geophys. Res. Lett., 36, doi:10.1029/2009gl040000, 2009a.

Huffman, G. J., Adler, R. F., Bolvin, D. T., and Nelkin, E. J.: The TRMM Multi-Satellite Precipitation Analysis (TMPA), in: Satellite Rainfall Applications for Surface Hydrology, pp. 3–22, Springer Netherlands, doi:10.1007/978-90-481-2915-7_1, 2009b.

Hundecha, Y., Arheimer, B., Donnelly, C., and Pechlivanidis, I.: A regional parameter estimation scheme for a pan-European multi-basin model, J. Hydrol. Regional Studies, 6, 90–111, doi:10.1016/j.ejrh.2016.04.002, 2016.

Legates, D. and Willmott, C.: Mean seasonal and spatial variability in gaugecorrected, global precipitation, Int. J. Climatol., 10, 111–127, 1990.

Li, H., Luo, L., Wood, E. F., and Schaake, J.: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting, J. Geophys. Res., 114(D4), D04 114, doi:10.1029/2008JD010969, 2009.

Lindström, G., Pers, C., Rosberg, R., Strömqvist, J., and Arheimer, B.: Development and test of the HYPE (Hydrological Predictions for the Environment) model - A water quality model for different spatial scales, Hydrol. Res., 41.3-4, 295–319, doi:10.2166/nh.2010.007, 2010.

Meier, M. F. and Bahr, D. B.: Counting glaciers: Use of scaling methods to estimate the number and size distribution of the glaciers of the world, in: Glaciers, Ice Sheets and Volcanoes: A Tribute to Mark F. Meier, vol. 96, pp. 89–94, DTIC Document, 1996.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andreassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 2 - Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, J. Hydrol., 303(1-4), 290–306, 2005.

Paiva, R. C. D., Collischonn, W., Bonnet, M. P., and De Goncalves, L. G. G.: On the sources of hydrological prediction uncertainty in the Amazon, Hydrol. Earth Syst. Sci., 16(9), 3127–3137, doi:10.5194/hess-16-3127-2012, 2012.

Pechlivanidis, I. G., Bosshard, T., Spångmyr, H., L. G., Gustafsson, D., and Arheimer, B.: Uncertainty in the Swedish operational hydrological forecasting systems, in: Vulnerability, Uncertainty, and Risk: Quantification, Mitigation and Management, edited by Beer, M., Au, S. K., and Hall, J. M., pp. 253–262, Liverpool, UK, doi:10.1061/9780784413609.026, 2014.

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Schneider, U.: GPCC First Guess Daily Product at 1.0°: Near Real-Time First Guess daily Land-Surface Precipitation from Rain-Gauges based on SYNOP Data, doi:10.5676/DWD_GPCC/FG_D_100, 2013.

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

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC Monitoring Product: Near Real-Time Monthly Land-Surface Precipitation from Rain-Gauges based on SYNOP and CLIMAT data, doi:10.5676/DWD_GPCC/MP_M_V5_100, 2015a.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data, doi:10.5676/DWD_GPCC/FD_M_V7_050, 2015b.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, J. Climate, 19, 3088–3111, 2006.

Shukla, S. and Lettenmaier, D. P.: Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill, Hydrol. Earth Syst. Sci., 15(11), 3529–3538, doi:10.5194/hess-15-3529-2011, 2011.

Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., and Lindström, G.: Water and nutrient predictions in ungauged basins - Set-up and evaluation of a model at the national scale, Hydrol. Sci. J., 57(2), 229–247, 2012.

Weedon, G., Gomes, S., Viterbo, P., Shuttleworth, W., Blyth, E., Österle, H., Adam, 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. Hydrometeor, 12, 823–848, doi: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, 2014.

Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Schneider, U.: GPCC First Guess Product at 1.0°: Near Real-Time First Guess monthly Land-Surface Precipitation from Rain-Gauges based on SYNOP Data, doi:10.5676/DWD_GPCC/FG_M_100, 2011.