Received: 26 January 2017 | Accepted: 28 June 2017
DOI: 10.1002/hyp.11269

WILEY

RESEARCH ARTICLE

Gridded climate data products are an alternative to instrumental measurements as inputs to rainfall–runoff models

José L.J. Ledesma | Martyn N. Futter

Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences (SLU), PO Box 7050, 750 07 Uppsala, Sweden

Correspondence
José L. J. Ledesma, Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences (SLU), PO Box 7050, 750 07 Uppsala, Sweden.
Email: jose.ledesma@slu.se

Funding information
Nordic eScience Globalisation Initiative (NeGI) — NordForsk, Grant/Award Number: 74306; Svenska Forskningsrådet Formas, Grant/Award Number: 2015-1518

Abstract

Rainfall–runoff models are widely used to predict flows using observed (instrumental) time series of air temperature and precipitation as inputs. Poor model performance is often associated with difficulties in estimating catchment-scale meteorological variables from point observations. Readily available gridded climate products are an underutilized source of temperature and precipitation time series for rainfall–runoff modelling, which may overcome some of the performance issues associated with poor-quality instrumental data in small headwater monitoring catchments. Here we compare the performance of instrumental measured and E-OBS gridded temperature and precipitation time series as inputs in the rainfall–runoff models "PERSiST" and "HBV" for flow prediction in six small Swedish catchments. For both models and most catchments, the gridded data produced statistically better simulations than did those obtained using instrumental measurements. Despite the high correspondence between instrumental and gridded temperature, both temperature and precipitation were responsible for the difference. We conclude that (a) gridded climate products such as the E-OBS dataset could be more widely used as alternative input to rainfall–runoff models, even when instrumental measurements are available, and (b) the processing applied to gridded climate products appears to provide a more realistic approximation of small catchment-scale temperature and precipitation patterns needed for flow simulations. Further research on this issue is needed and encouraged.

KEYWORDS

catchment science, E-OBS, hydrological modelling, model error, model input data, model uncertainty, precipitation measurement, temperature measurement

1 | INTRODUCTION

Hydrological modelling is essential for understanding runoff generation and solute transport processes. Modelling is subject to various types of uncertainty due to errors in input and calibration data (e.g., measurement errors and representativeness in time and space), model structure errors (e.g., inadequate or incorrect representation of processes and simplifications), and model parameter errors (e.g., "effective" vs. "actual" values and representativeness) (Beven, 2006; Clark, Kavetski, & Fenicia, 2011; Engeland, Xu, & Gottschalk, 2005). These multiple sources of uncertainty are not easily separated, leading to complex error structures and challenging hydrological simulation.

Whether the purpose of modelling is process understanding as part of hypothesis testing (Clark et al., 2011; Ruiz-Pérez et al., 2016) or political or industrial decision-making (Ledesma, Köhler, & Futter, 2012; Olsson & Andersson, 2007), modellers seek to reproduce natural processes as "realistically" as possible while maximizing model performance and minimizing uncertainty (Savenije, 2009). Thus, all modelling exercises attempt to reduce individual error sources as much as possible so as to constrain the overall uncertainty. Errors in model input data are one of the a priori simple and known types of uncertainty (Beven, 2006; Kuczera, Renard, Thyer, & Kavetski, 2010; Renard, Kavetski, Kuczera, Thyer, & Franks, 2010). Yet these errors are difficult to identify and correct, potentially leading to poor model performance.

In rainfall–runoff models, input data typically include daily time series of air temperature and precipitation (Arnold, Srinivasan, Muttiah, & Williams, 1998; Fenicia, Savenije, Matgen, & Pfister, 2006; Futter et al., 2014; Kampf & Burges, 2007; Lindström, Pers, Rosberg, Strömqvist, & Arheimer, 2010). Point observations of temperature and precipitation made at meteorological stations located on-site or nearby the study site are often used as model inputs (e.g., Abebe, Ogden, & Pradhan, 2010; Bernal, Butturini, Riera, Vázquez, & Sabater, 2004; Crossman et al., 2016; Oni et al., 2016). Point observation type of data will be referred to as instrumental measurements hereafter.

Instrumental measurements of meteorological variables, especially precipitation, are of concern when it comes to stream flow simulations (te Linde, Aerts, Hurkmans, & Eberle, 2008). Precipitation measurements are known to be subject to several different error sources including aerodynamic, wetting, evaporation, splash in and out, and blowing and drifting snow factors leading to uncertainty in estimates of rainfall and snowfall amounts (Taskinen & Söderholm, 2016). Very local storm events and microclimatic variations within the study catchment can also be problematic for the representativeness of measured precipitation (Orlowsky & Seneviratne, 2014), as can otherwise be unaccounted for factors that bias precipitation estimates (Sælthun, 1996). This issue becomes more evident as catchment area increases (Vaze et al., 2011). Temperature measurement errors are usually smaller but can also arise from thermometer exposure and urbanization (Folland et al., 2001). Temperature measurements are often more spatially representative than are precipitation measurements as temperature is generally less variable, especially in flat regions (Orlowsky & Seneviratne, 2014). Both temperature and precipitation measurements can also be subject to bad observer practices and data processing (Wilby et al., 2017). Ultimately, poor-quality data can lead to disinformation and incorrect model conditioning and calibration (Beven & Westerberg, 2011). Hence, it is essential to assess the quality of the meteorological data used as model inputs (Wilby et al., 2017), a process that is time consuming and, to some extent, subjective.

Given the aforementioned problems associated with instrumental observations of temperature and precipitation, there is a need to explore alternative data sources for catchment-scale rainfall–runoff modelling. This is especially relevant at headwaters and small monitoring catchments that are widely used for studying and understanding fundamental hydrological processes and that usually rely on instrumental observations of weather parameters. Gridded estimates of weather parameters derived from actual meteorological observations are one potential alternative data source. These gridded datasets have been used for runoff simulations in regions where instrumental data are lacking providing reasonably good results in some cases (Hadjikakou et al., 2011; Lauri, Räsänen, & Kummu, 2014; Vu, Raghavan, & Liong, 2012), but not always (Roth & Lemann, 2016; Yang, Wang, Wang, Yu, & Xu, 2014). The European Climate Assessment & Dataset project in Europe (ECA&D, 2017) that provides the E-OBS dataset (Haylock et al., 2008) and the Daymet project in North America (Daymet, 2017) are amongst the products that offer modellers the possibility to freely and easily access long time series of daily gridded climate data in regions where instrumental data are widely available. The question that arises is whether these gridded products can, or should, be used instead of actual instrumental measurements from on-site meteorological stations as inputs in rainfall–runoff models, especially at the small monitoring catchments that rely on such on-site measurements.

In order to answer this question, we tested how E-OBS gridded climate data compared with instrumental measurements from on-site or nearby meteorological stations for flow simulation in six small to medium size forest and agricultural catchments distributed over Sweden. Suitability of the two data sources was assessed in terms of model performance based on the Nash–Sutcliffe (NS) statistic comparing modelled and observed flows as simulated by two independent widely used rainfall–runoff models: Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport (PERSiST; Futter et al., 2014) and *Hydrologiska Byråns Vattenbalansavdelning* (HBV; Bergström, 1976; Bergström, 1992; Sælthun, 1996; Seibert, 2002). Rather than comparing HBV and PERSiST results or obtaining the best possible model fits, the main objective of this paper was to compare model efficiencies obtained by using on-site instrumental measurements of meteorological data versus gridded climate data as model inputs for flow simulations. Previous exercises have investigated how different climate data sources, especially rainfall data, compare for rainfall–runoff simulations in medium size to large catchments (e.g., Essou, Arsenault, & Brissette, 2016; Photiadou, Weerts, & van den Hurk, 2011; te Linde et al., 2008; Vaze et al., 2011). The focus of those studies was on the different model performances provided by different gridded products or on the utility of distributed versus lumped representations of rainfall for simulating stream flow in large catchments. In contrast, the small size of the catchments used here implies that observations made at locations inside or near the catchment are more likely to be representative of conditions within the catchment area, therefore allowing for direct comparison of the suitability of single point measurements versus single grid cell products as input data in stream flow modelling. To our knowledge, this is the first time that the E-OBS dataset has been compared with instrumental measurements as forcing variables of rainfall–runoff models using a consistent calibration strategy.

2 | MATERIAL AND METHODS

Stream flow was simulated in six forest and agricultural Swedish catchments (Table 1) using the rainfall–runoff models PERSiST and HBV. Equal ranges were given to each of the parameters that are commonly sensitive in each model during a Monte Carlo approach to calibration. The process was carried out using, on the one hand, instrumental meteorological data and, on the other hand, the E-OBS gridded climate data as model inputs. A total of 24 different model calibrations were performed (6 catchments × 2 models × 2 sets of input data). Results were compared in terms of model efficiency based on the NS statistic (Nash & Sutcliffe, 1970).

2.1 | Study sites and flow data

Six small to medium size, well-studied Swedish catchments with a range in land use and climate were used (Table 1). All six catchments are located in areas of low relief. Their hydrographs (Figure S1) are

TABLE 1 Characteristics and basic information of the study catchments

Information/parameter	Gårdsjön	Kindla	Gammtratten	Svartberget	C6	Sävjaån
Area (km^2)	0.037	0.204	0.448	0.470	32.9	725
Latitude (N)	58°03′	59°05′	63°51′	64°15′	59°43′	59°50′
Longitude (E)	12°01′	14°54′	18°06′	19°47′	17°09′	17°40′
Elevation range (m)	114–140	312–415	420–540	235–310	10–59	1–72
Forest (%)	84	71	86	82	40	66
Wetland (%)	11	24	14	18	3	3
Agriculture (%)	0	0	0	0	57	31
Bedrock (%)	5	5	0	0	0	0
Calibration period	2006–2012	2006–2012	2006–2012	2006–2012	1996–2009	1996–2009
Distance weather station—Catchment outlet (km)	0.1	0.2	0.9	1.0	10	5.0
Mean annual measured precipitation (mm)	1171	833	653	657	562	561
Mean annual gridded precipitation (mm)	1088	860	680	640	545	575
Mean measured temperature (°C)	7.2	4.9	2.4	2.4	6.7	6.7
Mean gridded temperature (°C)	7.3	5.2	2.3	3.0	6.7	6.6
Daily precipitation R^2	.37	.74	.83	.54	.90	.92
Monthly precipitation R^2	.70	.65	.89	.92	.96	.94
Daily temperature R^2	.99	.99	.93	.99	.99	.99
Monthly temperature R^2	.99	.99	.95	.99	.99	.99

Note. Coefficients of determination (R^2) were obtained from regressions of instrumental versus gridded climate data at daily and monthly time scales. These relationships were statistically significant at $p < .001$.

characterized by intra-annual variability with snowmelt (more pronounced in northern sites) and summer–autumn rainfall episodes (more pronounced in southwestern sites).

Three of the sites (Gårdsjön, Kindla, and Gammtratten) are part of the Swedish integrated monitoring catchments (IM sites; Löfgren et al., 2011; Lundin et al., 2001). A fourth site, commonly known as Svartberget (Bishop, Grip, & O′Neill, 1990), is part of the Krycklan Catchment Study, an intensively studied infrastructure for experimental and hypothesis-driven research in the boreal landscape (Laudon et al., 2013). These four sites are all small (3.7 to 47 ha) forest-dominated headwater catchments. They are all relatively undisturbed and cover a climate gradient across Sweden. The granitic and gneissic bedrock is overlain in all cases by Quaternary deposits of glacial till. Forests stands of Norway spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) dominate in upslope podzols, whereas riparian zones are organic-rich soils (histosols) with no pines and small proportions of birch (*Betula* spp.). Stream discharge at Gårdsjön and Svartberget was measured using 90° V-notch weirs. Established stage–discharge rating curves were used to transform the registered water levels at the weirs into stream flow. At Kindla and Gammtratten, discharge was recorded using H-flumes, which have slightly different geometry but follow the same principle as V-notch weirs. In these four catchments, stream discharge data were available for 2006–2012.

Two other larger, predominantly agricultural catchments (C6 and Sävjaån) were included so as to have a wider range of size and land cover (Table 1). Catchment C6 (33 km^2) is one of the 21 catchments that are monitored to study nutrient losses from agricultural land within the Swedish Environmental monitoring program (Kyllmar, Carlsson, Gustafson, Ulén, & Johnsson, 2006). Discharge at C6 was measured using a 90° V-notch weir. Sävjaån (725 km^2) is part of a long-term routine monitoring program conducted by the Department of Aquatic Science and Assessment (*Institutionen for Vatten och Miljö*) at the Swedish University of Agricultural Sciences (Fölster, Johnson, Futter, & Wilander, 2014). Discharge at Sävjaån was measured at the river outlet using a water level device in a section of known dimensions. Available stream discharge data for these two catchments covered the 14-year period 1996–2009.

2.2 | Instrumental meteorological data

At the IM sites, air temperature is measured using on-site sensors, and precipitation is measured using tipping buckets. Data collection, processing, and handling of meteorological variables measured at these sites follow the manual for quality assurance prepared by the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (Raspe, Beuker, Preuhsler, & Bastrup-Birk, 2016). Daily values were used for calibration periods (2006–2012). A reference climate monitoring program that follows the World Meteorological Organization recommendations was established at the Krycklan Catchment Study in 1980 (Laudon et al., 2013) and was used as source of daily temperature and precipitation data for the Svartberget catchment calibration (2006–2012).

Daily temperature and precipitation for the calibration period (1996–2009) for C6 and Sävjaån were obtained from the Swedish Meteorological and Hydrological Institute (*Sveriges Meteorologiska och Hydrologiska Institut* [SMHI]) meteorological stations located in the cities of Uppsala (a few kilometres northwest of the Sävjaån catchment outlet) and Enköping (approximately 10 km south of the C6 catchment outlet). Temperature was not available from the station at Enköping, so temperature data from Uppsala were used for catchment C6 (separated by about 30 km).

All instrumental measurement data presented here have been previously used as model inputs in a variety of hydrological and water quality modelling efforts (Exbrayat et al., 2010; Futter et al., 2011;

Karlsen et al., 2016; Kyllmar et al., 2006). Thus, they are considered representative of actual conditions at the different catchments, especially at the smaller sites where temperature and precipitation are measured on-site and weather variability within the catchments is small.

2.3 | E-OBS gridded climate data

The European high-resolution gridded dataset "E-OBS" consists of daily values of precipitation and minimum, maximum, and mean surface temperature back to 1950 for 0.25 × 0.25° grids over Europe (Haylock et al., 2008). Grid size approximately spans from 330 to 420 km^2 at the study catchment latitudes. The dataset was developed as part of the EU-FP6 project ENSEMBLES (2017) and allows direct comparison with Regional Climate Models (RCMs).

Raw temperature and precipitation observations from contributing meteorological stations undergo a series of quality tests to identify data issues (Haylock et al., 2008; Hofstra, Haylock, New, & Jones, 2009). This allows correction or removal of suspicious values. Subsequently, instrumental observations are spatially and temporally interpolated in a three-step process to provide best estimates of grid box averages. Interpolations are performed separately for temperature and precipitation (Haylock et al., 2008; Hofstra et al., 2009). For temperature, monthly mean values were estimated using stations with <20% missing data for that month. These monthly means were then interpolated using thin-plate splines to represent the underlying spatial trends. Daily anomalies with respect to monthly means were kriged, and the resultant kriging estimator was applied to the monthly anomaly to generate a final result (Haylock et al., 2008). Daily gridded precipitation estimates were generated as follows. First, monthly means were estimated on the basis of stations with <20% missing data for that month. Indicator kriging was then performed to identify days on which precipitation was assumed to fall, on the basis of a threshold of 0.5 mm. Daily gridded precipitation estimates were then generated on the basis of whether precipitation was assumed to be falling and the fraction of the total monthly precipitation falling on that day (Haylock et al., 2008).

The number of contributing stations reported by Haylock et al. (2008) was 2316, and the updated website list as for November 2016 showed around 11000, of which over 1500 are located in Sweden. The uncertainty of the spatial interpolation is larger when the number of contributing stations is lower (Haylock et al., 2008); for example, northern Sweden is a region with limited station coverage (Hofstra et al., 2009). The data can be freely accessed and downloaded at the ECA&D website (http://eca.knmi.nl/download/ensembles/download.php#datafiles). Daily mean temperature and daily precipitation from the grid cells where the study catchments are located were used in the calibrations presented here. As the Sävjaån catchment area is bigger than a single grid cell, the location of the catchment outlet was used to select the appropriate grid cell because it is also the closest point to the meteorological station used to obtain the instrumental data.

2.4 | Rainfall–runoff model characterizations

A brief description of the two rainfall–runoff models applied is given below.

2.4.1 | Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport

PERSiST (Futter et al., 2014) is a semidistributed, bucket-type model for daily flow simulations. It consists of a flexible framework that allows the modeller to specify the perceptual representation of the runoff generation process, which is based on a number of interconnected buckets within a mosaic of landscape units in the basin. PERSiST requires daily time series of air temperature and precipitation as input data.

Rainfall and snowmelt are directed to the stream as overland flow or infiltrated to the soil, which is divided into a number of specified layers. Depending on the bucket structure, soil water can move vertically to lower soil layers or return to upper layers, the soil surface, or the atmosphere, or move horizontally downhill or to the stream. Water movement is controlled by field capacities, hydrological connectivities, and infiltration-related parameters. Snowfall is deposited and accumulated on the ground. Snowmelt and water lost via evapotranspiration are controlled by degree day rates and threshold temperatures. Precipitation as rain or snow can also be intercepted by the canopy. The magnitude and flashiness of the simulated flow are also dependent on the catchment area and water velocity-related parameters.

2.4.2 | Hydrologiska Byråns Vattenbalansavdelning

HBV, developed by SMHI (Bergström, 1976; Bergström, 1992), is a semidistributed conceptual rainfall–runoff model that has been widely used for flow simulations and forecasting (Abebe et al., 2010; Seibert, McDonnell, & Woodsmith, 2010; Steele-Dunne et al., 2008). HBV simulates daily flow by using daily time series of temperature and precipitation as input data. Potential evaporation of the 12 months was estimated using the Thornthwaite equation and added as another necessary input to HBV light (Seibert, 2002; Seibert & Vis, 2012), a user-friendly version of HBV with a sophisticated Monte Carlo routine.

Runoff generation process representation in HBV is similar to that in PERSiST, including a snow routine, a soil moisture routine, a response function, and a routing routine. The more flexible representation of terrestrial hydrology and model structure in PERSiST allows, a priori, simulating a wider range of hydrologic conditions.

2.5 | Model calibration strategy

For each catchment, model calibration periods were set according to the available stream flow data (Table 1). The strategy that was followed to calibrate the 24 model instances (6 catchments × 2 models × 2 sets of input data) was analogous in all cases so as to have consistent results that could be compared in terms of model efficiency. Typical sensitive parameters in PERSiST (Futter et al., 2014; Oni et al., 2016) were assigned common ranges (Table 2) and set in a Monte Carlo approach to model calibration (Steele-Dunne et al., 2008). In each case, a total of 100,000 model runs were executed and the best 100 of those, in terms of model efficiency based on the NS statistic, were kept for comparison. Analogously, sensitive HBV parameters (Abebe et al., 2010) were given common ranges (Table 2) within the HBV light version. The

TABLE 2 Parameter ranges during model calibrations

Parameter name and small description	Units	Min	Max
	PERSiST		
a: Flow velocity multiplier	–	0.01	0.2
b: Flow velocity exponent	–	0.6	0.95
Snow multiplier	–	0.75	1.3
Rain multiplier	–	0.75	1.3
Snow melt temperature	°C	−2	2
Degree day melt factor	mm °C^{-1}	0.5	4
Degree day evapotranspiration	mm °C^{-1}	0.05	0.3
Growing degree threshold	°C	−1	3
Time constant quick box	day	1	2
Time constant fast box	day	1.5	6
Time constant slow box	day	5	20
Drought runoff fraction fast box	–	0	0.2
	HBV		
TT: Threshold temperature for snow simulation	°C	−2	2
CFMAX: Degree day factor	mm °C^{-1}	0.5	4
SFCF: Snowfall correction factor	–	0.5	0.9
FC: Maximum soil moisture storage	mm	100	500
LP: Soil moisture above which Et_{act} reaches Et_{pot}	mm	0.3	1
BETA: Relative contribution to runoff from precipitation	–	1	7
PERC: Maximum percolation rate	mm day^{-1}	0	4
UZL: Threshold for lateral flow movement	mm	0	100
K0: Recession coefficient upper box	day^{-1}	0.1	0.9
K1: Recession coefficient middle box	day^{-1}	0.01	0.2
K2: Recession coefficient lower box	day^{-1}	0.001	0.002
MAXBAS: Length of triangular weighting function	day	1	7

Note. HBV = Hydrologiska Byråns Vattenbalansavdelning; PERSiST = Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport.

total number of model runs during each Monte Carlo calibration was also 100,000, and the 100 highest NS efficiencies were kept for comparison in all cases. Importantly, the same number of parameters (12 in total) was varied in both models to have the same number of degrees of freedom and thus the same degree of potential overfitting.

2.6 | Statistical analyses

A three-way factorial analysis of variance using the best 100 NS of all 24 calibrations was performed using the statistical software JMP 13.0 to estimate whether gridded climate data produced different model performance as compared with instrumental data. Site and model were set as random effects, whereas the type of input data was set as a fixed effect.

Statistical comparisons based on Tukey's honestly significant difference tests were performed for each individual catchment for the NS efficiencies of the 100 best-performing parameter of the four corresponding calibrations (PERSiST or HBV and gridded or instrumental). The following pairs were compared: PERSiST gridded–PERSiST instrumental; HBV gridded–HBV instrumental; PERSiST gridded–HBV gridded; and PERSiST instrumental–HBV instrumental.

3 | RESULTS

3.1 | Instrumental versus E-OBS gridded climate data

Differences between measured and gridded mean annual temperature and annual precipitation were low for calibration periods, except for temperature at Svartberget where the gridded annual was 0.6°C higher than instrumental values (Table 1). Hereafter, overestimations or underestimations refer to E-OBS gridded data as compared with instrumental measurements. Subannual comparisons still showed a good correspondence between measured and gridded temperature in most cases (data not shown), except for the northern catchments where there was an underestimation in 2011 for Gammtratten and a general overestimation in winter at Svartberget (responsible for the mean annual discrepancy).

Precipitation on the other hand showed some mismatches that were irrespective of season and had no clear pattern. For example, at Gårdsjön, a southern catchment, there were large overestimations in summer rainfall in 2006 and 2007 (192 and 72 mm higher), but large underestimations in 2009, 2010, and 2012 (87 to 160 mm lower). There was also a general overestimation in autumn rainfall in Gammtratten and summer rainfall in Kindla. At monthly and daily scales, temperature patterns were very similar (R^2 > .93) between gridded and instrumental data in all cases (Table 1). Precipitation comparison plots were more scattered, especially the daily regressions at Gårdsjön and Svartberget as compared to the good correspondence at C6 and Sävjaån (Figure 1).

3.2 | Model result comparisons

Model NS efficiencies of the 100 best-performing parameter sets (Figure 2) were compared for each individual catchment using Tukey's honestly significant difference tests (Table 3). Hereafter, better/higher or worse/lower implies a statistically significant difference between NS produced by a particular calibration in comparison to another.

There was only one case in which instrumental data produced a better calibration than did the E-OBS gridded data and that was the PERSiST application in Svartberget. The E-OBS data produced better performances than did instrumental data using both models in the IM sites (Gårdsjön, Kindla, and Gammtratten), for the PERSiST application in C6, and for the HBV application in Sävjaån. Model calibrations obtained by HBV were better than were those obtained by PERSiST in Sävjaån, whereas PERSiST provided higher NS efficiencies using both gridded and instrumental data for the rest of the comparisons except for Gårdsjön that showed no difference between models (Table 3).

The analysis of variance (Table 4) showed that, overall, the E-OBS gridded data produced an NS performance of 0.058 units higher than did the instrumental data, or about a 10% increase. Similarly, PERSiST produced an NS performance of 0.055 units higher than did HBV, which corresponded to a 9% increase.

4 | DISCUSSION

The E-OBS gridded climate dataset was tested as an alternative to instrumental measurements of temperature and precipitation as inputs in rainfall–runoff models. We showed that gridded data produced better simulations of stream flow than did those obtained using

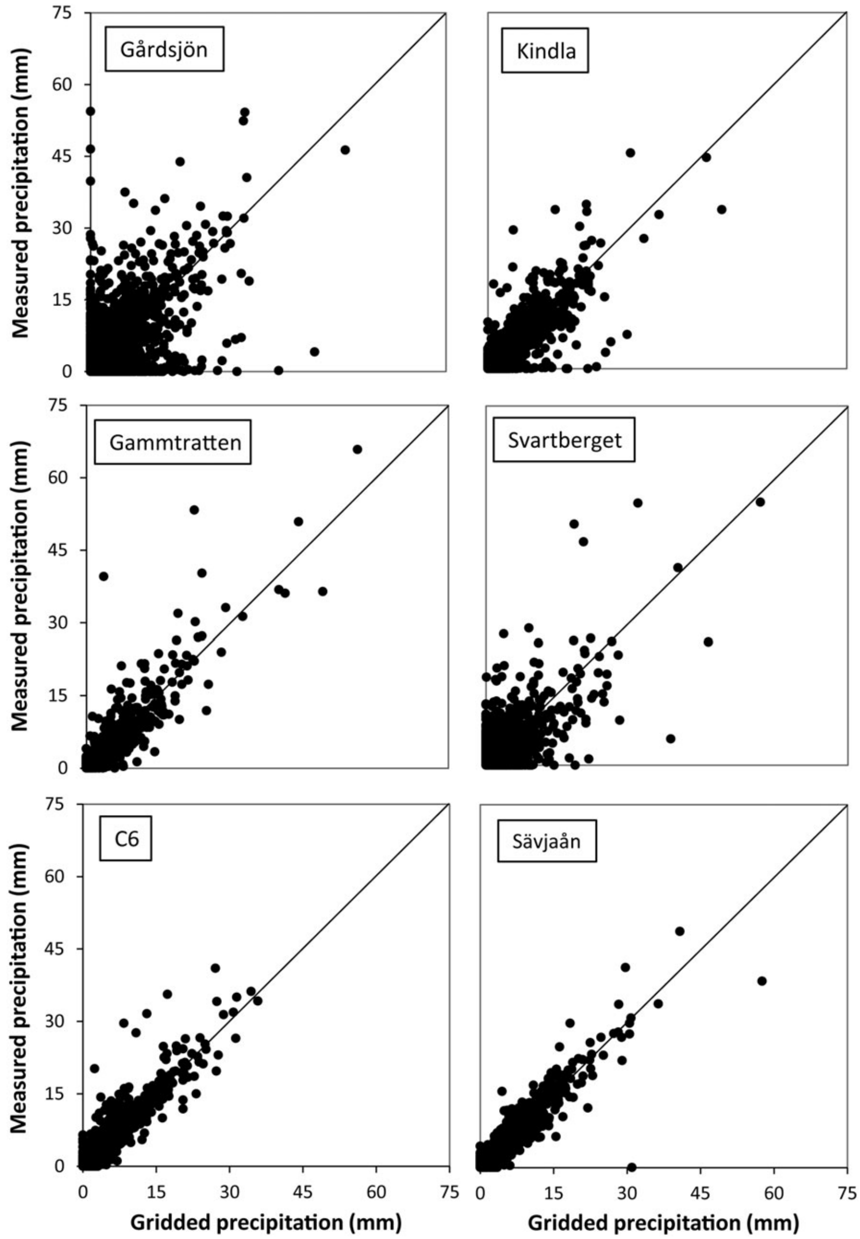

FIGURE 1 Regression plots of instrumental measured versus gridded daily precipitation for the study sites

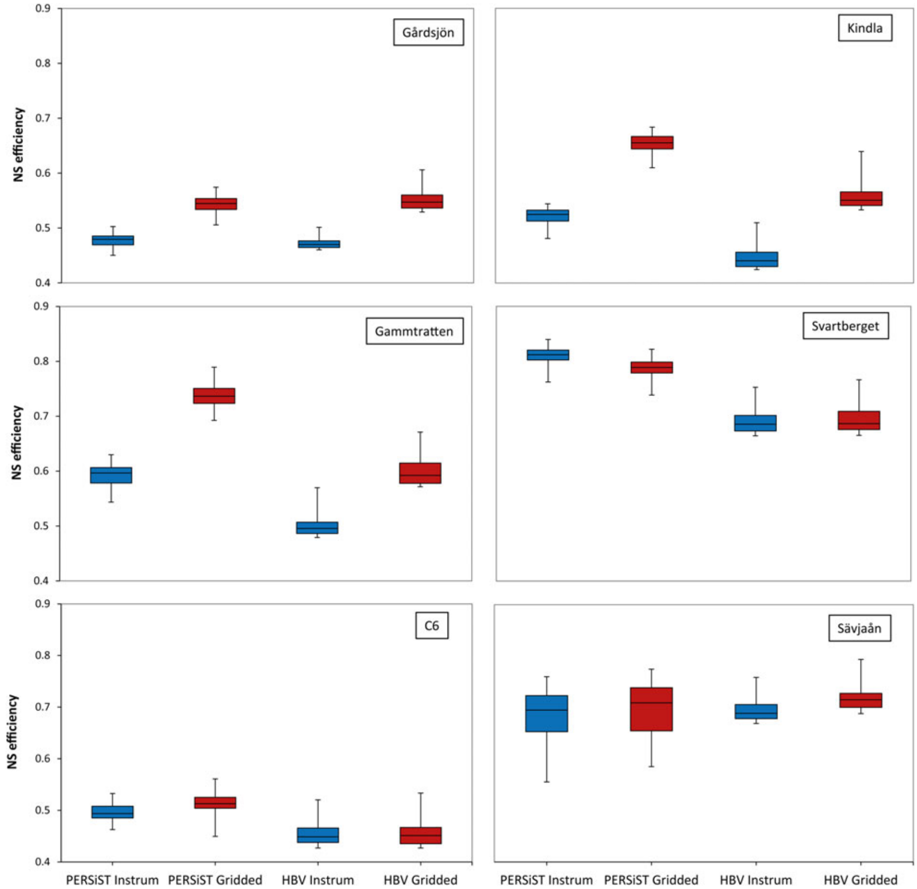

FIGURE 2 Box plots of model calibration efficiencies based on the 100 best Nash-Sutcliffe (NS) performances including six catchments, two models (PERSiST or HBV), and two sets of input data (instrumental in blue or gridded in red). HBV = Hydrologiska Byråns Vattenbalansavdelning; PERSiST = Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport

TABLE 3 Tukey's honestly significant difference test for the Nash-Sutcliffe (NS) efficiencies of the 100 best-performing parameter sets in the specified calibration comparisons

Comparison	Gårdsjön	Kindla	Gammtratten	Svartberget	C6	Sävjaån
PERSiST gridded–PERSiST measured	Gridded***	Gridded***	Gridded***	Instrumental***	Gridded***	ns
HBV gridded–HBV measured	Gridded***	Gridded***	Gridded***	ns	ns	Gridded***
PERSiST gridded–HBV gridded	ns	PERSiST***	PERSiST***	PERSiST***	PERSiST***	HBV***
PERSiST measured–HBV measured	ns	PERSiST***	PERSiST***	PERSiST***	PERSiST***	HBV***

Note. The specified dataset or model (either gridded or instrumental, or PERSiST or HBV) was the one providing a better performance.

HBV = Hydrologiska Byråns Vattenbalansavdelning; ns = not significant; PERSiST = Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport.

***$p < .0001$.

instrumental meteorological observations. By using a set of different catchment sizes, land use type proportions, climatic conditions, sources and methods for instrumental data, and two independent models, we minimized potential bias related to any of these factors. We also argue that the potential for model overfitting, a common issue in overparameterized models (Beven, 2006), is not relevant here because the same number of parameters were allowed to vary during all calibrations, leading to the same degree of fitting across all model runs regardless of model or dataset used. Thus, the results of this exercise were interpreted in terms of the relative difference in model efficiencies and, consequently, in the suitability of input data for stream flow simulations.

Instrumental and E-OBS gridded temperature had, in general, a high degree of correspondence. Therefore differences in model efficiencies could, a priori, be attributed mainly to the discrepancies in precipitation between instrumental and gridded data. To test this further, we reran both PERSiST and HBV at all sites using the combination of E-OBS gridded precipitation and instrumental temperature as inputs. Model performances obtained with the combined gridded and instrumental data for the sites with small previous differences (Svartberget, C6, and Sävjaån) were very similar to those obtained with either only gridded or only instrumental data (Figure S2). Interestingly, although the combined dataset for Gårdsjön gave very similar results to those obtained with only gridded data, it produced intermediate

TABLE 4 Three-way factorial analysis of variance using the best 100 Nash-Sutcliffe (NS) of all 24 model calibrations

Source	*df*	Sum of squares	*F* ratio	Prob > *F*
Site	5	20.8	7596	<.0001
Data	1	2.0	3624	<.0001
Model	1	1.8	3261	<.0001
Site × Data	5	1.9	692	<.0001
Site × Model	5	1.8	644	<.0001
Data × Model	1	0.02	33.9	<.0001
Site × Data × Model	5	0.14	50.8	<.0001

Note. Site and model (PERSiST or HBV) were set as random effects, whereas the type of data (gridded or instrumental) was set as a fixed effect. HBV = Hydrologiska Byråns Vattenbalansavdelning; PERSiST = Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport.

performances in the case of Gammtratten and Kindla (Figure S2). Thus, as expected, precipitation was mainly responsible for the better performance of gridded data at Gårdsjön, but a combination of temperature and precipitation was responsible for the better performance of gridded data at Gammtratten and Kindla. In contrast to previous assessments (Photiadou et al., 2011) and studies where only rainfall data sources were considered to test differences in model performances (Vaze et al., 2011; Vu et al., 2012), our results suggest that different temperature data sources could also be responsible for those differences and should be taken into account. This is important even if, as in here, the degree of correspondence between temperature datasets is high.

In any case, it is intriguing that in all cases but one, gridded data could be used to simulate stream flow as well as or better than instrumental measured data, thereby providing a good alternative when instrumental measurements are lacking or problematic. Photiadou et al. (2011) showed that a gridded dataset constructed by the Commission for the Hydrology of the Rhine base (CHR) used as input data in HBV outperformed the E-OBS dataset in a series of large catchments in the Rhine basin. This CHR dataset was an extended version of that used by te Linde et al. (2008), who also showed that the CHR produced better performances than data from an RCM in two rainfall–runoff models, including HBV. This, a priori, indicates that gridded products based on observations provide better model performances than did datasets from RCMs. It also indicates that specific gridded datasets might outperform the E-OBS dataset. However, a limitation of the approach presented by Photiadou et al. (2011) and te Linde et al. (2008) was that HBV was only calibrated using the CHR data and then forced with the other time series for the comparisons.

Here, the E-OBS data, which are based on real observations and not climate model outputs, were used in model calibration in an analogous method as the instrumental measurements. Measured data used in the E-OBS product are subjected to a systematic quality check so as to identify issues and to correct or remove nonsensible values (Haylock et al., 2008). The process of interpolation includes thin-plate spline (Hutchinson, 1995) and kriging (Atkinson & Lloyd, 1998), which homogenizes data both spatially and temporally. These methods aim to achieve a good spatial representation but also produce data time series that are structurally consistent minimizing inhomogeneity, variance, and randomness. This usually reduces the magnitude of extremes, which, a priori, is not desirable in rainfall–runoff models that need those extremes to fit hydrological events. Hofstra et al. (2009) warn that the use of the E-OBS dataset, which underestimates extremes of precipitation, may cause an underestimation of high flows. However, the E-OBS data here reproduced the hydrological extremes as well as or better than did the instrumental data as indicated by the NS statistic, which tends to be biased toward fitting high flows (Jain & Sudheer, 2008). Therefore, it appears that rainfall–runoff model fits are favoured by internally consistent input data time series such as spatially interpolated gridded products as compared to point observation time series such as on-site instrumental measurements. This could in fact be sensible as, for example, any consistent bias in the gridded data could be corrected by model parameters such as rain multiplier (an adjustment factor to relate input data to the actual rainfall at the site used in PERSiST), BETA (relative contribution to runoff from precipitation used in HBV), growing degree threshold (temperature above which evapotranspiration can occur used in PERSiST), or snow melt temperatures (temperature above which snow melts used in both PERSiST and HBV). Similar versions of these parameters are also common in other widely used rainfall–runoff models (Arnold et al., 1998; Lindström et al., 2010).

The methods for measuring temperature and, especially, precipitation, likely also influenced the results. Precipitation at the SMHI stations used as source of instrumental data for the catchments C6 and Sävjaån was measured by automatic weighing gauges, which is described as a sound method, more reliable than tipping buckets (Sevruk, Ondrás, & Chvíla, 2009). These stations are in fact listed as contributing stations to the E-OBS program. Still, the efficiencies of the models using the E-OBS data were higher than those using instrumental SMHI data in the mesoscale catchments C6 and Sävjaån (Table 3; Figure 2). This further supports that gridded processed, internally consistent time series are good-quality inputs in rainfall–runoff models even in places where on-site or nearby measurements are robust. Differences in the quality control of the raw data might have also played a role. At Svartberget, where high-standard methodologies following the World Meteorological Organization recommendations are used to record temperature and precipitation (Laudon & Ottosson-Löfvenius, 2016; Laudon et al., 2013), the gridded dataset provided worse fits than did instrumental measurements when using PERSiST, but equally good fits when using HBV. It is therefore still possible that high-quality instrumental meteorological data measured on-site outperform gridded products in small catchments. However, this was only true for one of the two models used in Svartberget. Similarly, the differences found in Svartberget, C6, and Sävjaån, even if statistically significant in some cases due to the high statistical power, are not as obvious as for the smaller sites Gårdsjön, Kindla, and Gammtratten (Figure 2) and it could be argued that gridded and instrumental data perform equally as well. Tipping buckets, used to record precipitation in those three smaller sites, likely provides an example of situations when gridded data could indeed substitute for instrumental measurements as these three catchments showed remarkably better model fits with the E-OBS dataset.

Although the purpose of the exercise was not to compare performance of the two models, it could not go unnoticed that PERSiST outperformed HBV in most cases. Exploring the reasons for this is outside the scope of this paper, but it appears that the more flexible representation of terrestrial hydrology in PERSiST might help to simulate a wider range of hydrologic conditions (Futter et al., 2014). The goal of the paper was not to achieve the best possible model fits either. The parameter space was not exhaustively explored in any of the calibrations, and the full capabilities of PERSiST and HBV were not employed. For example, PERSiST allows the user to alter the routing of water within the soil, but this was not considered and a fixed proportion was set in all soil layers and for all catchments. Variations of this option would have been helpful to obtain better fits at the smaller, flashier catchments. This may partially account for the relatively low model efficiencies obtained in some of the calibrations.

5 | CONCLUSIONS

We argue that gridded climate data products, such as the E-OBS data, can be a viable alternative to instrumental data as inputs to rainfall–runoff models even in well-instrumented regions such as Europe. This principle potentially applies also to North America, as suggested by Essou et al. (2016) who compared different gridded datasets for hydrological modelling including the Daymet dataset (Thornton, Running, & White, 1997), and Australia, as suggested by Vaze et al. (2011) who compared instrumental rainfall data and the SILO gridded rainfall data (Jeffrey, Carter, Moodie, & Beswick, 2001) across the country for rainfall–runoff calibration and simulation. Nevertheless, those studies focused on significantly larger catchments than the ones presented here so we argue that gridded climate data can be also an alternative in small (well-studied or not) sites. This alternative should be considered when on-site meteorological measurements are lacking, incomplete, unavailable, suspicious, erroneous, inconsistent, or, importantly, costly, as gridded data appear to be a safe and easy way forward to set up and calibrate rainfall–runoff models. It should be noted that the quality of the gridded product is related to the number of contributing meteorological stations (Haylock et al., 2008). Thus, gridded data in poorly instrumented regions might be useful but do not appear to perform as well as instrumental data (Vu et al., 2012). A further advantage of, for example, the E-OBS dataset is the long consistent period of record back to 1950, thereby having a high potential to hind cast or calibrate and simulate past stream flows.

The E-OBS data could then be used as an alternative to actual instrumental measurements when the methods to record meteorological parameters, especially precipitation, are not optimal or have known issues, or when their values need to be regularly corrected. Monitoring of environmental parameters including meteorological variables is essential to maintain and develop catchment science (Fölster et al., 2014; Lovett et al., 2007). However, there is a current trend to reduce environmental funding in some parts of the world, which may imply the need to adjust economic budgets in, for example, research field stations in monitoring catchments. If this is the case, weather stations that record meteorological variables in small research catchments could be prioritized to be expended, as freely accessible gridded climate products could provide a reliable alternative to obtain this type of data. Many scientists who are not experts in the mathematical algorithms of spatial interpolation are still in the need of using climate data such as the gridded data presented here or those obtained from RCMs. Consumers of such data include catchment scientists, biogeochemists, and ecologists, who should be aware of the existence of gridded dataset products that can be a reliable material for their modelling purposes. Yet a broader test of, in this case, the E-OBS dataset, including more catchments and more locations within Europe, could prove useful for supporting our conclusions. The use of gridded products as an alternative to instrumental measurements is likely to become even more feasible as additional high-resolution products are developed (e.g., Prein et al., 2016).

A second important implication of this paper is that rainfall–runoff models seem to work better in terms of fitting flow observations when input data time series have internal patterns of coherent variability, that is, less noise and lower inhomogeneity and variance. This is the case for spatially interpolated gridded data. The interpolation algorithm would have a decisive role on the final outcome (Vu et al., 2012), but this opens new questions about our current understanding of hydrological models. We suggest that the processing applied to gridded climate products can provide a more realistic approximation of small catchment-scale temperature and precipitation patterns than that obtained from point observations, and this could be the reason why the gridded data produced better flow simulations than did the instrumental data. Further research that provides expanded answers on this issue is necessary and encouraged.

ACKNOWLEDGMENTS

We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). We thank Lars Lundin and Stefan Löfgren for providing the IM sites data and helpful discussions and Joachim Audet for guidance on the C6 catchment data. Finally, we thank two anonymous referees for their reviews, which helped to improve the quality of the paper.

REFERENCES

Abebe, N. A., Ogden, F. L., & Pradhan, N. R. (2010). Sensitivity and uncertainty analysis of the conceptual HBV rainfall–runoff model: Implications for parameter estimation. *Journal of Hydrology*, *389*, 301–310. https://doi.org/10.1016/j.jhydrol.2010.06.007

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment—Part 1: Model development. *Journal of the American Water Resources Association*, *34*, 73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x

Atkinson, P. M., & Lloyd, C. D. (1998). Mapping precipitation in Switzerland with ordinary and indicator kriging. Special issue: Spatial interpolation comparison 97. *Journal of Geographic Information and Decision Analysis*, *2*, 72–86.

Bergström, S. (1976). Development and application of a conceptual runoff model for Scandinavian catchments. *SMHI RHO*, 7, Norrköping, 134 pp.

Bergström, S. (1992). The HBV model: Its structure and applications. *SMHI RH*, 4, Norrköping, 35 pp.

Bernal, S., Butturini, A., Riera, J. L., Vázquez, E., & Sabater, F. (2004). Calibration of the INCA model in a Mediterranean forested catchment: The effect of hydrological inter-annual variability in an intermittent stream. *Hydrology and Earth System Sciences*, *8*, 729–741.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*, 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Beven, K., & Westerberg, I. (2011). On red herrings and real herrings: Disinformation and information in hydrological inference. *Hydrological Processes*, *25*, 1676–1680. https://doi.org/10.1002/hyp.7963

Bishop, K. H., Grip, H., & O′Neill, A. (1990). The origins of acid runoff in a hillslope during storm events. *Journal of Hydrology*, *116*, 35–61. https://doi.org/10.1016/0022-1694(90)90114-d

Clark, M. P., Kavetski, D., & Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, *47*. W09301. doi: https://doi.org/10.29/2010wr009827

Crossman, J., Eimers, M. C., Watmough, S. A., Futter, M. N., Kerr, J., Baker, S. R., & Dillon, P. J. (2016). Can recovery from disturbance explain observed declines in total phosphorus in Precambrian shield catchments? *Canadian Journal of Fisheries and Aquatic Sciences*, *73*, 1202–1212. https://doi.org/10.1139/cjfas-2015-0312

Daymet. https://daymet.ornl.gov/, last access: 9 January 2017.

ECA&D (European Climate Assessment & Dataset). http://eca.knmi.nl/, last access: 9 January 2017.

Engeland, K., Xu, C. Y., & Gottschalk, L. (2005). Assessing uncertainties in a conceptual water balance model using Bayesian methodology. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *50*, 45–63. https://doi.org/10.1623/hysj.50.1.45.56334

ENSEMBLES: http://ensembles-eu.metoffice.com/, last access: 9 January 2017.

Essou, G. R. C., Arsenault, R., & Brissette, F. P. (2016). Comparison of climate datasets for lumped hydrological modeling over the continental United States. *Journal of Hydrology*, *537*, 334–345. https://doi.org/10.1016/j.jhydrol.2016.03.063

Exbrayat, J. F., Viney, N. R., Seibert, J., Wrede, S., Frede, H. G., & Breuer, L. (2010). Ensemble modelling of nitrogen fluxes: Data fusion for a Swedish meso-scale catchment. *Hydrology and Earth System Sciences*, *14*, 2383–2397. https://doi.org/10.5194/hess-14-2383-2010

Fenicia, F., Savenije, H. H. G., Matgen, P., & Pfister, L. (2006). Is the groundwater reservoir linear? Learning from data in hydrological modelling. *Hydrology and Earth System Sciences*, *10*, 139–150.

Folland, C. K., Rayner, N. A., Brown, S. J., Smith, T. M., Shen, S. S. P., Parker, D. E., … Sexton, D. M. H. (2001). Global temperature change and its uncertainties since 1861. *Geophysical Research Letters*, *28*, 2621–2624. https://doi.org/10.1029/2001gl012877

Fölster, J., Johnson, R. K., Futter, M. N., & Wilander, A. (2014). The Swedish monitoring of surface waters: 50 years of adaptive monitoring. *Ambio*, *43*, 3–18. https://doi.org/10.1007/s13280-014-0558-z

Futter, M. N., Erlandsson, M. A., Butterfield, D., Whitehead, P. G., Oni, S. K., & Wade, A. J. (2014). PERSiST: A flexible rainfall–runoff modelling toolkit for use with the INCA family of models. *Hydrology and Earth System Sciences*, *18*, 855–873. https://doi.org/10.5194/hess-18-855-2014

Futter, M. N., Löfgren, S., Köhler, S. J., Lundin, L., Moldan, F., & Bringmark, L. (2011). Simulating dissolved organic carbon dynamics at the Swedish integrated monitoring sites with the integrated catchments model for carbon, INCA-C. *Ambio*, *40*, 906–919. https://doi.org/10.1007/s13280-011-0203-z

Hadjikakou, M., Whitehead, P. G., Jin, L., Futter, M., Hadjinicolaou, P., & Shahgedanova, M. (2011). Modelling nitrogen in the Yeşilirmak River catchment in northern Turkey: Impacts of future climate and environmental change and implications for nutrient management. *Science of the Total Environment*, *409*, 2404–2418. https://doi.org/10.1016/j.scitotenv.2011.02.038, 2011

Haylock, M. R., Hofstra, N., Tank, A., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research-Atmospheres*, *113*. D20119, doi: https://doi.org/10.1029/2008jd010201

Hofstra, N., Haylock, M., New, M., & Jones, P. D. (2009). Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. *Journal of Geophysical Research-Atmospheres*, *114*. D21101, doi: https://doi.org/10.1029/2009jd011799

Hutchinson, M. F. (1995). Interpolating mean rainfall using thin-plate smoothing splines. *International Journal of Geographical Information Systems*, *9*, 385–403. https://doi.org/10.1080/02693799508902045

Jain, S. K., & Sudheer, K. P. (2008). Fitting of hydrologic models: A close look at the Nash–Sutcliffe index. *Journal of Hydrologic Engineering*, *13*, 981–986. https://doi.org/10.1061/(asce)1084-0699(2008)13:10(981)

Jeffrey, S. J., Carter, J. O., Moodie, K. B., & Beswick, A. R. (2001). Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling & Software*, *16*, 309–330. https://doi.org/10.1016/s1364-8152(01)00008-1

Kampf, S. K., & Burges, S. J. (2007). A framework for classifying and comparing distributed hillslope and catchment hydrologic models. *Water Resources Research*, *43*. W05423, doi: https://doi.org/10.1029/2006wr005370

Karlsen, R. H., Grabs, T., Bishop, K., Buffam, I., Laudon, H., & Seibert, J. (2016). Landscape controls on spatiotemporal discharge variability in a boreal catchment. *Water Resources Research*, *52*, 6541–6556. https://doi.org/10.1002/2016wr019186

Kuczera, G., Renard, B., Thyer, M., & Kavetski, D. (2010). There are no hydrological monsters, just models and observations with large uncertainties! *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *55*, 980–991. https://doi.org/10.1080/02626667.2010.504677

Kyllmar, K., Carlsson, C., Gustafson, A., Ulén, B., & Johnsson, H. (2006). Nutrient discharge from small agricultural catchments in Sweden characterisation and trends. *Agriculture Ecosystems & Environment*, *115*, 15–26. https://doi.org/10.1016/j.agee.2005.12.004

Laudon, H., & Ottosson-Löfvenius, M. O. (2016). Adding snow to the picture—Providing complementary winter precipitation data to the Krycklan catchment study database. *Hydrological Processes*, *30*, 2413–2416. https://doi.org/10.1002/hyp.10753

Laudon, H., Taberman, I., Ågren, A., Futter, M., Ottosson-Löfvenius, M., & Bishop, K. (2013). The Krycklan catchment study—A flagship infrastructure for hydrology, biogeochemistry, and climate research in the boreal landscape. *Water Resources Research*, *49*, 7154–7158. https://doi.org/10.1002/wrcr.20520

Lauri, H., Räsänen, T. A., & Kummu, M. (2014). Using reanalysis and remotely sensed temperature and precipitation data for hydrological modeling in monsoon climate: Mekong River case study. *Journal of Hydrometeorology*, *15*, 1532–1545. https://doi.org/10.1175/jhm-d-13-084.1

Ledesma, J. L. J., Köhler, S. J., & Futter, M. N. (2012). Long-term dynamics of dissolved organic carbon: Implications for drinking water supply. *Science of the Total Environment*, *432*, 1–11. https://doi.org/10.1016/j.scitotenv.2012.05.071

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. (2010). Development and testing of the HYPE (hydrological predictions for the environment) water quality model for different spatial scales. *Hydrology Research*, *41*, 295–319. https://doi.org/10.2166/nh.2010.007

Löfgren, S., Aastrup, M., Bringmark, L., Hultberg, H., Lewin-Pihlblad, L., Lundin, L., … Thunholm, B. (2011). Recovery of soil water, groundwater, and streamwater from acidification at the Swedish integrated monitoring catchments. *Ambio*, *40*, 836–856. https://doi.org/10.1007/s13280-011-0207-8

Lovett, G. M., Burns, D. A., Driscoll, C. T., Jenkins, J. C., Mitchell, M. J., Rustad, L., … Haeuber, R. (2007). Who needs environmental monitoring? *Frontiers in Ecology and the Environment*, *5*, 253–260. https://doi.org/10.1890/1540-9295(2007)5[253:wnem]2.0.co;2

Lundin, L., Aastrup, M., Bringmark, L., Bråkenhielm, S., Hultberg, H., Johansson, K., … Löfgren, S. (2001). Impacts from deposition on Swedish forest ecosystems identified by integrated monitoring. *Water Air and Soil Pollution*, *130*, 1031–1036. https://doi.org/10.1023/a:1013956228299

Nash, J. E., & Sutcliffe, V. (1970). River flow forecasting through conceptual models: Part I—A discussion of principles. *Journal of Hydrology*, *10*, 282–290.

Olsson, J. A., & Andersson, L. (2007). Possibilities and problems with the use of models as a communication tool in water resource management.

Water Resources Management, *21*, 97–110. https://doi.org/10.1007/s11269-006-9043-1

Oni, S., Futter, M., Ledesma, J., Teutschbein, C., Buttle, J., & Laudon, H. (2016). Using dry and wet year hydroclimatic extremes to guide future hydrologic projections. *Hydrology and Earth System Sciences*, *20*, 2811–2825. https://doi.org/10.5194/hess-20-2811-2016

Orlowsky, B., & Seneviratne, S. I. (2014). On the spatial representativeness of temporal dynamics at European weather stations. *International Journal of Climatology*, *34*, 3154–3160. https://doi.org/10.1002/joc.3903

Photiadou, C. S., Weerts, A. H., & van den Hurk, B. (2011). Evaluation of two precipitation data sets for the Rhine River using streamflow simulations. *Hydrology and Earth System Sciences*, *15*, 3355–3366. https://doi.org/10.5194/hess-15-3355-2011

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Maule, C. F., van Meijgaard, E., Déque, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., & Jacob, D. (2016). Precipitation in the EURO-CORDEX 0.11 degrees and 0.44 degrees simulations: High resolution, high benefits? *Climate Dynamics*, *46*, 383–412, doi: https://doi.org/10.1007/s00382-015-2589-y, 2016.

Raspe, S., Beuker, E., Preuhsler, T., & Bastrup-Birk, A. (2016). Part IX: Meteorological measurements. In: UNECE ICP Forests Programme Co-ordinating Centre. In *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests* (pp. 35). Eberswalde, Germany: Thünen Institute of Forest Ecosystems. http://www.icpforests.org/Manual.htm

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, *46*. W05521, doi: https://doi.org/10.1029/2009wr008328

Roth, V., & Lemann, T. (2016). Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT in small catchments in the Ethiopian Highlands. *Hydrology and Earth System Sciences*, *20*, 921–934. https://doi.org/10.5194/hess-20-921-2016

Ruiz-Pérez, G., Medici, C., Latron, J., Llorens, P., Gallart, F., & Francés, F. (2016). Investigating the behaviour of a small Mediterranean catchment using three different hydrological models as hypotheses. *Hydrological Processes*, *30*, 2050–2062. https://doi.org/10.1002/hyp.10738

Sælthun, N. R. (1996). The Nordic HBV model. *Norwegian Water Resources and Energy Administration Publication*, *7*, 1–26.

Savenije, H. H. G. (2009). HESS Opinions "The art of hydrology". *Hydrology and Earth System Sciences*, *13*, 157–161.

Seibert, J. (2002). HBV light version 2, User's manual. *Department of Environmental Assessment, SLU*, Uppsala, 32 pp.

Seibert, J., McDonnell, J. J., & Woodsmith, R. D. (2010). Effects of wildfire on catchment runoff response: A modelling approach to detect changes in snow-dominated forested catchments. *Hydrology Research*, *41*, 378–390. https://doi.org/10.2166/nh.2010.036

Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*, 3315–3325. https://doi.org/10.5194/hess-16-3315-2012

Sevruk, B., Ondrás, M., & Chvíla, B. (2009). The WMO precipitation measurement intercomparisons. *Atmospheric Research*, *92*, 376–380. https://doi.org/10.1016/j.atmosres.2009.01.016

Steele-Dunne, S., Lynch, P., McGrath, R., Semmler, T., Wang, S. Y., Hanafin, J., & Nolan, P. (2008). The impacts of climate change on hydrology in Ireland. *Journal of Hydrology*, *356*, 28–45. https://doi.org/10.1016/j.jhydrol.2008.03.025

Taskinen, A., & Söderholm, K. (2016). Operational correction of daily precipitation measurements in Finland. *Boreal Environment Research*, *21*, 1–24.

te Linde, A. H., Aerts, J., Hurkmans, R., & Eberle, M. (2008). Comparing model performance of two rainfall–runoff models in the Rhine basin using different atmospheric forcing data sets. *Hydrology and Earth System Sciences*, *12*, 943–957.

Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, *190*, 214–251. https://doi.org/10.1016/s0022-1694(96)03128-9

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Teng, J., & Viney, N. R. (2011). Conceptual rainfall–runoff model performance with different spatial rainfall inputs. *Journal of Hydrometeorology*, *12*, 1100–1112. https://doi.org/10.1175/2011jhm1340.1

Vu, M. T., Raghavan, S. V., & Liong, S. Y. (2012). SWAT use of gridded observations for simulating runoff—A Vietnam river basin study. *Hydrology and Earth System Sciences*, *16*, 2801–2811. https://doi.org/10.5194/hess-16-2801-2012

Wilby, R. L., Clifford, N. J., De Luca, P., Harrigan, S., Hillier, J. K., Hodgkins, R., … Wood, P. J. (2017). The 'dirty dozen' of freshwater science: Detecting then reconciling hydrological data biases and errors. *Wiley Interdisciplinary Reviews—Water*, *4*, 1–19. https://doi.org/10.1002/wat2.1209

Yang, Y., Wang, G. Q., Wang, L. J., Yu, J. S., & Xu, Z. X. (2014). Evaluation of gridded precipitation data for driving SWAT model in area upstream of Three Gorges reservoir. *PloS One*, *9*, 15. https://doi.org/10.1371/journal.pone.0112725

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ledesma JLJ, Futter MN. Gridded climate data products are an alternative to instrumental measurements as inputs to rainfall–runoff models. *Hydrological Processes*. 2017;1–11. https://doi.org/10.1002/hyp.11269