Constraining the HBV model for robust water balance assessments in a cold climate

Helene Birkelund Erlandsen, Stein Beldring, Stephanie Eisner, Hege Hisdal, Shaochun Huang and Lena Merete Tallaksen

ABSTRACT

Robust projections of changes in the hydrological cycle in a non-stationary climate rely on trustworthy estimates of the water balance elements. Additional drivers than precipitation and temperature, namely wind, radiation, and humidity are known to have a significant influence on processes such as evaporation, snow accumulation, and snow-melt. A gridded version of the rainfall-runoff HBV model is run at a 1×1 km scale for mainland Norway for the period 1980–2014, with the following alterations: (i) the implementation of a physically based evaporation scheme; (ii) a net radiation-restricted degree-day factor for snow-melt, and (iii) a diagnostic precipitation phase threshold based on temperature and humidity. The combination of improved forcing data and model alterations allowed for a regional calibration with fewer calibrated parameters. Concurrently, modeled discharge showed equally good or better validation results than previous gridded model versions constructed for the same domain; and discharge trend patterns, snow water equivalent, and potential evaporation compared fairly to observations. Compared with previous studies, lower precipitation and evaporation values for mainland Norway were found. The results suggest that a more robust and more physically based model for climate change studies has been obtained, although additional studies will be needed to further constrain evaporation estimates.

Key words | climate, evaporation, HBV, Norway, snow-melt

Helene Birkelund Erlandsen (corresponding author)
Stein Beldring
Hege Hisdal
Shaochun Huang
Norwegian Water Resources and Energy Directorate,
Oslo,
Norway
E-mail: *helenebe@met.no*

Helene Birkelund Erlandsen
Lena Merete Tallaksen
University of Oslo,
Oslo,
Norway

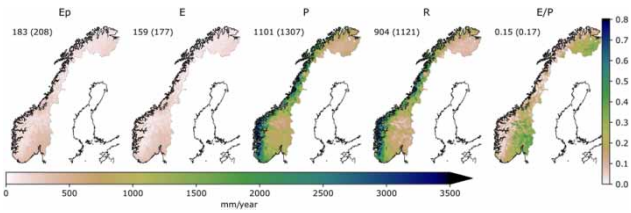
Stephanie Eisner
Norwegian Institute of Bioeconomy Research (NIBIO),
Ås,
Norway

HIGHLIGHTS

- The distributed HBV model is updated with physically based parameterizations.
- High-quality forcing data are included to enhance estimates evaporation, precipitation phase, and snow-melt.
- More than 100 discharge measurements are used for calibration and validation.
- The updates help constrain the long-term water balance for Norway.
- Additional work is called for to better constrain the evaporation estimates.

doi: 10.2166/nh.2021.132

GRAPHICAL ABSTRACT



INTRODUCTION

Many hydrological models were developed for operational water resources management and have accordingly been built to rely only on input data that is commonly available, and to be easy to use. Any increase in model complexity should be justified by an increase in model performance, often measured according to the ability of the model to reproduce daily or monthly catchment runoff (Nash & Sutcliffe 1970; Lindström *et al.* 1997; Ferguson 1999). Today, gridded input data are becoming more widely available, either from numerical weather prediction models, reanalysis data such as Era5 (Hersbach *et al.* 2018), gridded observational data, e.g. SeNorge2018 (Lussana *et al.* 2019), or hybrid products such as HySN (Erlandsen *et al.* 2019) and WFDEI (Weedon *et al.* 2014). Further, the non-stationarity of the current climate calls for hydrological models with a stronger physical basis and a higher robustness in a wide range of climates (Ferguson 1999; Clark *et al.* 2015).

Hydrological models range from the simplest, data-driven, lumped, and conceptually based water balance models, to those akin to land surface models, where the surface energy balance is solved numerically (see e.g. Kauffeldt *et al.* 2016). The different modeling strategies have complimenting merits (Hrachowitz & Clark 2017). For example, numerically solving the surface energy balance requires an increase in input data requiring higher storage and pre-processing capacity, as well as an increase in model integration time; however, it allows the computation of surface temperature, and imposing a closed surface energy balance, which further constrains the latent heat flux or evaporation estimates. In this paper, the term evaporation encompasses water loss from soil, leaves, lakes, and plant stomata (transpiration).

A large number of hydrological modeling studies involve replacing a rather simple conceptual process description with a more physically based one and comparing the results (Bruland *et al.* 2001; Zappa *et al.* 2003; Hegdahl *et al.* 2016), while other studies compare models of different complexities (e.g. Magnusson *et al.* 2015). In cases where a more physically based model was compared with a more conceptually based model and an increase in model performance was not found, it is difficult to say whether this was due to an ill-stated empirical equation or parameter being included, over-parameterization, or that the more physically based process description relied on input variables that were poorly estimated. Thus, an important question raised in Clark *et al.* (2015) is 'to what extent is additional model complexity supported by the available information on geophysical attributes (topography, vegetation, soils, geology, and fine-scale meteorological data)?'.

A particular challenge for using conceptual, calibrated models for hydrological impact assessment is that parameter values can be overfitted to the climate conditions in the calibration period. Merz *et al.* (2011) found that calibrated parameters representing snow and soil moisture processes were sensitive to the choice of the calibration period. Milly & Dunne (2011) found that a temperature-index based evaporation parameterization may simulate considerably larger evaporation changes than net radiation changes might justify. Any change in vapor pressure deficit with climate change, or plant physiological mechanisms for preserving water, are also not accounted for in temperature-index based evaporation parameterizations. Besides evaporation-related calibrated parameters, other often calibrated parameters which potentially might be omitted

from calibration, or considerably restricted in range, are the precipitation phase threshold temperature and the commonly used degree-day factors for calculating snow-melt. Precipitation phase may be diagnosed using near-surface temperature and humidity (Jennings *et al.* 2018). The degree-day factor, which represents the amount of snow-melt per degree above freezing, may vary considerably depending on catchment, climate, and time-of-year (Kustas *et al.* 1994; Merz *et al.* 2011). Indirectly, it reflects biases in accumulated snowfall, sublimation and deposition processes not accounted for, spatial and temporal variation in incident longwave and shortwave radiation, vegetation shading, longwave radiation emitted by vegetation, and variation in surface albedo, to name a few. Accordingly, in climate change studies, for some processes in particular, there is a need to move from a simple and conceptually based description to a more robust, physically based one.

In Norway, a gridded version of the conceptual HBV rainfall-runoff model (Beldring *et al.* 2003, from here on referred to as HBV-B03) has been used to study the effect of climate change on hydrology (see e.g. Hanssen-Bauer *et al.* 2017). Until recently HBV-B03 included calibrated, land cover-dependent parameters for precipitation phase diagnosis, the melting temperature of snow, the snow-melt degree-day factor, and for the temperature-based scaling of monthly climatological potential evaporation to provide estimates of evaporation. In Wong *et al.* (2011), the existing evaporation routine was discussed as a large source of uncertainty when analyzing end-of-century changes in summer droughts for Norway. Further, the lack of an in-line computation of potential evaporation may be particularly unsuitable in cold climates – since potential evaporation is limited by the received incident radiation, which is bound to increase in a warmer climate with reduced snow cover and thus albedo.

There has been a recent effort to improve the physical basis of evaporation estimates in HBV-B03 by implementing a Penman–Monteith (Monteith 1965) potential evaporation routine. Simultaneously, the number of land cover classes represented by the model was increased from 7 to 19 to allow for more spatial heterogeneity related to natural vegetation cover and land use activity. These alterations are described in Huang *et al.* (2019), from here on this version of the model is referred to as HBV-H19. The inclusion of a more detailed land cover description combined with land use dependent, calibrated parameters controlling processes such as snow accumulation and ablation may lead to confounded or poorly constrained parameters, and thus disentanglement problems if the model were to be applied to study the effect of a perturbed land cover. Kustas *et al.* (1994) suggested an enhanced degree-day factor parameterization, where the degree-day factor is restricted by an additive term relating snow-melt to net radiation. A snow-melt routine where the degree-day factor is restricted by a radiative term allows snow-melt to be influenced by land cover class via albedo, without the need of a land cover class-dependent calibration.

The implementation of more physically based process descriptions as described above is here facilitated by a newly established hybrid method, HySN, for producing gridded estimates of near-surface vapor pressure and incident radiation. HySN was derived by merging reanalysis data with the 1×1 km SeNorge data and showed high fidelity when compared with station observations (Erlandsen *et al.* 2019).

In this study, we aim at obtaining a robust and more physically based model for studies of changes in water balance elements in a non-stationary climate. The HySN method for deriving estimates of evaporation and incident radiation was paired with an improved version of the SeNorge temperature and precipitation fields, SeNorge2018 (Lussana *et al.* 2019) and used as forcing data for a modified version of HBV-H19. The availability of high-quality input data made way for adding the following physically based updates to the HBV model:

i. an augmented Penman–Monteith based evaporation scheme;
ii. a regionally calibrated, radiation-restricted degree-day factor;
iii. a diagnostic temperature- and humidity-based threshold for diagnosing precipitation phase.

Simulated discharge was evaluated in terms of bias and Kling–Gupta Efficiency (KGE; Gupta *et al.* 2009) using measurements from more than 100 catchments, of which 34 were independent, i.e. not used for calibration. Calibration was conducted for the period 2000–2010. In addition, an independent validation time period, 1980–1999,

was identified. The model's suitability for climate studies was assessed by evaluating its ability to reproduce monthly discharge trends, simulated maximum winter snow water equivalent (SWE), and by comparing estimated potential evaporation with pan evaporation measurements. The model was run from 1980 to 2014. Its simulated mean water balance was assessed and compared with previous water balance estimates for Norway.

STUDY AREA AND DATA

Study area

Mainland Norway stretches several latitudes, from 58° to 71° North, on the western coast of northern Europe (see Figure 1). Its coastline is lined with fjords, while further inland the Scandinavian Mountains divide the country's western and eastern regions. Norway's location on the eastern end of the North Atlantic and its prevailing westerly winds, together with the Scandinavian Mountains, leads to a high annual precipitation on its western coast, with distinctly lower precipitation rates received leeward of the mountain range. Around a third of precipitation falls as snow. About 38% of the land area is forest covered, while the land surface is dominated by bare rock and shallow deposits (see e.g. Figure 17.1, replication of the Geological Survey of Norway (NGU) sediment map in Weynants *et al.* (2013)). Relatively shallow soils with a low water storage capacity in large parts of the country make way for a rapid runoff response to precipitation (e.g. Beldring 2002), but also for moisture stressed conditions in periods of meteorological drought (Buckland *et al.* 1997).

Forcing data

The model is forced with gridded daily temperature and precipitation fields from SeNorge2018, and with surface incident shortwave radiation, surface net longwave radiation, and vapor pressure deficit derived following the HySN method, and with a 1×1 km resolution wind data set from the Norwegian Meteorological Institute (MET Norway). Details of the forcing data are given below.

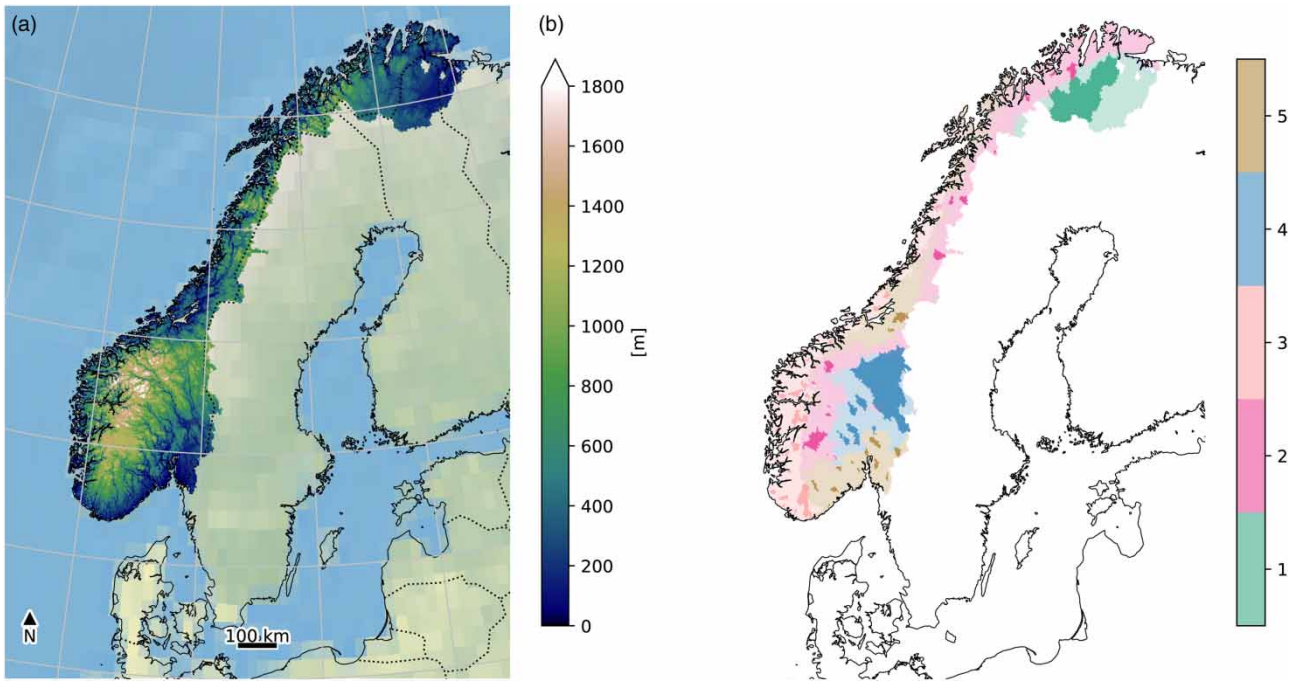


Figure 1 | The study region, Norway, with catchment areas outside Norway which drain into Norway included. (a) The orography and (b) the five regions used in the model calibration, with catchments used for calibration shown with a more opaque color. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/nh.2021.132.

SeNorge2018

SeNorge2018 version 18.12 is the newest version of the 1 × 1 km gridded data sets of 2-meter temperature (T2) and precipitation (P) based on observations from surface meteorological stations, developed by MET Norway. The data have a daily resolution and cover the period 1957 until the present. SeNorge2018 includes several innovations compared with the previous SeNorge version, such as the inclusion of a wind-induced undercatch correction for precipitation based on Wolff *et al.* (2015), and the use of climatological background fields from a convection-permitting dynamical downscaling of the global reanalysis ERA-Interim (Dee *et al.* 2011) instead of observational gridded data. For the model integration period, 1980–2014, SeNorge2018 shows an annual mean precipitation of 1,348 mm for Norway, while the previous version, SeNorge2.1 (Lussana *et al.* 2018), which did not include precipitation undercatch corrections, shows 1,068 mm, i.e. 280 mm less (see Supplementary Figure S1).

Klinogrid wind

The wind data set used to force HBV is a high-resolution, quantile-mapping-based gridded data set of near-surface wind speed developed at MET Norway. The daily wind data is available for October 1957 until May 2015 from http://thredds.met.no/thredds/catalog/metusers/klinogrid/KliNoGrid_16.12/FFMRR-Nor/catalog.html (accessed 13 December 2019).

HySN5

HySN5 is a modified version of HySN, a high-resolution HYbrid SeNorge data set of daily near-surface humidity, surface incident shortwave and longwave radiation, and surface pressure. The data have the same temporal frequency and projection as the SeNorge data sets. It is described and compared with surface observations and other data sets in Erlandsen *et al.* (2019). The downscaling procedure used to produce HySN5 is unchanged from Erlandsen *et al.* (2019); however, Era5 has replaced Era-Interim and SeNorge2018 has replaced SeNorge2. The assumptions and methods used to downscale humidity and longwave radiation are similar to those used in the WATCH and WFDEI data sets (Weedon *et al.* 2014), PGMFD (Sheffield *et al.* 2006), and NLDAS-1 (Cosgrove 2003).

HySN5 is compared with surface observations for the same time period as used in Erlandsen *et al.* (2019) (1982–1999), including 84 stations where measurements of 2-meter humidity are available, and 10 (2) stations where incident shortwave (longwave) radiation are observed. The comparison shows that HySN5 vapor pressure has a similar mean daily correlation with station measurements, 0.95, and a slightly lower mean absolute station bias than HySN, 31 kPa rather than 35 kPa (see Supplementary Figure S2). The incident shortwave radiation of HySN5 shows a slightly higher daily correlation with observations (0.95) than HySN (0.94, see Supplementary Figure S3); however, it also shows a larger mean difference to the station observations (7.9 W m^{-2}) than HySN (3.2 W m^{-2}). A proper validation of incident longwave radiation is difficult since only two stations are available, both situated near the coast; nevertheless, at the two stations, HySN5 shows a higher daily correlation with observations than HySN, 0.94 rather than 0.91; however, also here HySN5 shows a higher average difference to the observations (Supplementary Figure S4).

Vapor pressure deficit was estimated by calculating vapor pressure at saturation using SeNorge2018 T2. Net longwave radiation was calculated assuming a surface emissivity of 0.96 and that surface temperature can be approximated by T2. These two derived variables are available on request. At present, HySN5 is available from Zenodo from 1979 through 2000 (https://doi.org/10.5281/zenodo.3351430) and from 2001 through 2017 (https://doi.org/10.5281/zenodo.3516560).

Calibration and validation data

Discharge observations, quality controlled and available from The Norwegian Water Resources and Energy Directorate (NVE), for 119 catchments across Norway, were used for calibration and validation. Additional evaluation data includes 24,148 observations from 1,181 measurement sites of SWE. SWE was derived from snow depth and

density, which are routinely measured, predominantly measured in the mountainous regions of southern Norway, by hydropower companies (Saloranta 2012), usually once a year, around the time of maximum SWE.

Few measurements of evaporation are available for Norway as compared with other Nordic countries. Pan evaporation was measured in the summer season, May through September, between 1967 and 1972. Table 7.5 in Hetager & Lystad (1974) lists the mean monthly Pan evaporation, covering 3–5 years, for 42 stations (reproduced in the Supplementary Material). The measurements were retrieved by a Thorsrud 2500 evaporimeter, composed of a 50 cm deep sunken pan with a diameter of 56 cm, filled with water.

METHODS

The gridded HBV model

HBV-B03

HBV-B03 (Beldring *et al.* 2003) has a daily resolution and covers mainland Norway with 1 km^2 grid cells. The model performs water balance calculations for square grid cells characterized by their elevation, land use, and soil type. Each grid cell includes glacial and lake fractions, and up to three land cover classes, and one soil type. It has a snow routine with components for accumulation, and sub-grid scale distribution and ablation of snow. It also includes glacier melt. The evaporation routine includes a land cover-dependent evaporation, lake evaporation, and interception storage. Soil moisture and discharge are simulated including a sub-grid scale distribution of soil moisture storage, groundwater storage, and runoff response. Six soil parameters are routinely calibrated: field capacity (FC), an exponent controlling the fraction of infiltration that percolates to the upper zone (beta), a parameter controlling the percolation to the lower zone (perc), the upper (kuz) and lower (klz) zones' runoff response coefficients, and the upper zone recession parameter (alpha). The model further includes a routing module; however, routing has not been included in either of the model implementations discussed in this study.

HBV-H19

In HBV-H19 (Huang *et al.* 2019), HBV-B03 was modified by replacing its simple, temperature-based scaling of potential evaporation (E_p) with a Penman–Monteith big leaf approximation for estimating daily E_p. E_p parameters were set as fixed parameters based on the literature or physical empirical relationships in a look-up table; however, maximum interception storage was for most land cover classes calibrated according to the region and land cover class. The precipitation phase threshold temperature was set to 0 °C, while the threshold temperature for snow-melt and the degree-day factor were calibrated according to the region and land cover class. The minimum and maximum degree--day factor allowed during calibration was 0.0001 and 0.01 m $°C^{-1}$, respectively (Table 1 in Huang *et al.* (2019)). The land cover class was diagnosed based on the high-resolution National Land Resource Map (Ahlström *et al.* 2014) combined with a structural forest classification map provided by Majasalmi *et al.* (2018). Forests were classified into three species groups (spruce, pine, and deciduous) with four structure classes, each reflecting a low (class 1) to high (class 4) biomass density, among other attributes. The number of land cover and soil types represented by the model was increased to 19 and 12, respectively (see Huang *et al.* 2019). The model domain was divided into five calibration regions (see Figure 1(b)), which resulted from *k*-means clustering of two temperature and two precipitation indices.

HBV-E20

The model version used in this study, HBV-E20, builds upon HBV-B03 and HBV-H19. The land cover and soil type classification are based on the same data sets as in HBV-H19; however, the resulting classification is slightly modified for land cover, while the soil type classification was simplified, allowing a total of seven classes for all of Norway. Figure 2 depicts the dominant land cover class and soil type in each grid cell.

The evaporation scheme is similar to that described in HBV-H19. The Supplementary Material provides an overview of the implemented Penman–Monteith algorithm (Equation S3). In HBV-E20, surface resistance (r_s) is

Table 1 | Summary statistics comparing model simulated discharge and observations for the calibrated catchments in the calibration period and the independent period, and for independent basins

	Number of catchments	KGE (–)	Mean bias (simulated – observed) (mm/day)	Mean observed discharge (mm/day)	Relative bias (%)
Calibration period (2000–2010)	79	0.71 (0.74)	−0.3	17.2	−1.9
Independent period (1980–1999)	76	0.70 (0.75)	−0.1	13.6	−0.7
Independent basins (1980–2010)	34	0.74 (0.78)	−0.8	17.0	−4.8

modified from HBV-H19. The new approach describes r_s as a function of humidity deficit, visible radiation at the top of the canopy, and leaf area index, according to Leuning *et al.* (2008). Standard parameters are used, as recommended by Leuning *et al.* (2008), with maximum stomatal conductance given as a function of land cover class, and based on literature recommendations (Schulze *et al.* 1994; Kelliher *et al.* 1995; Körner 1995). Additionally, the surface resistance formula in Leuning *et al.* (2008) is modified to include a temperature constraint, as described in Mu *et al.* (2011). Look-up table values with parameters related to the 19 land cover types are provided in Supplementary Table S3.

HBV-E20 further includes changes to the model's snow-melt routine. We implement a radiation-restricted degree-day factor based on Kustas *et al.* (1994). A radiation-based melt rate, in meters per day, obtained by converting the R_n to snow-melt rate, is added to the common degree-day factor expression:

$$M = \max\left(C_{\text{temp}}(T2 - T2_{\text{melt}}) + C_{\text{rad}}\frac{R_n}{\lambda_f \rho_w}, 0\right) \quad (1)$$

where M is the melt rate per day in meters, C_{temp} is a calibrated degree-day factor, and $T2_{\text{melt}}$ is the melt temperature of snow, λ_f is the latent heat of fusion, 0.334 MJ kg^{-1}, R_n is

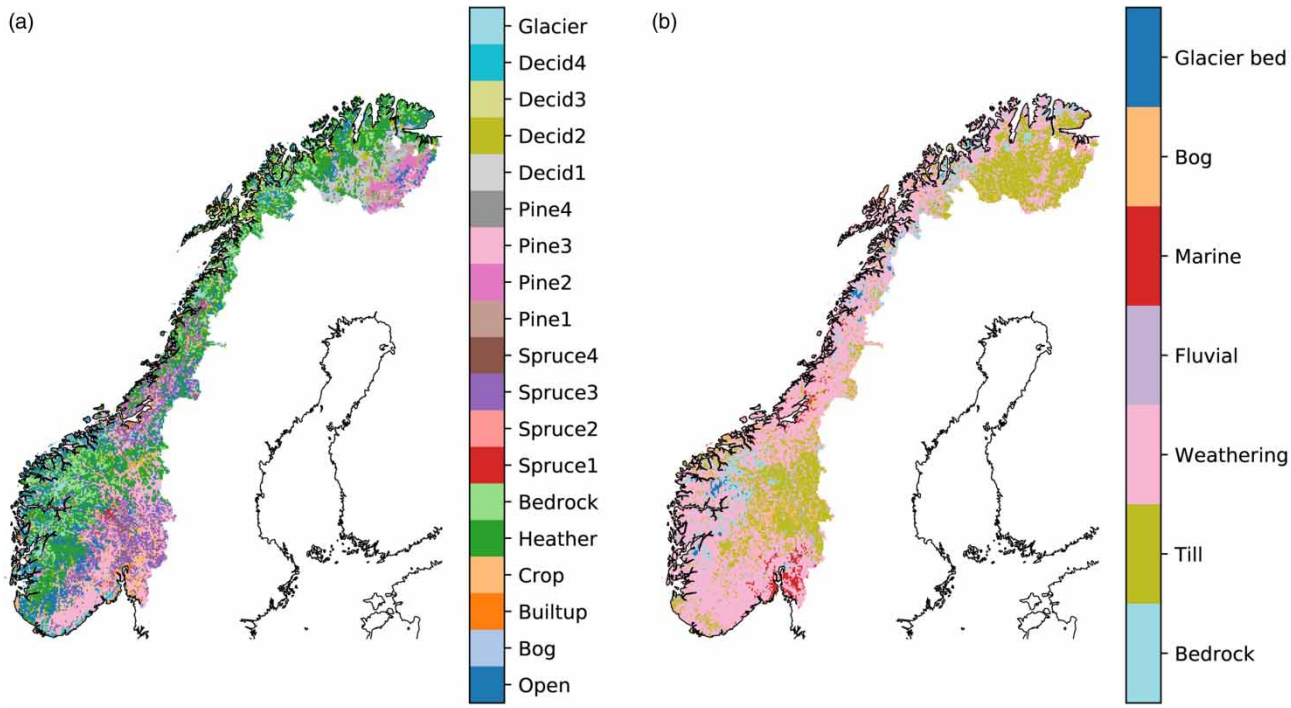


Figure 2 | The dominant land cover class within each grid cell is shown in (a), while (b) shows the soil type. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/nh.2021.132.

in MJ m^{-2} day^{-1}, ρ_w is the density of water (1,000 kg m^{-3}), and C_{rad} is a fraction scaling the radiation term, which is always less than unity (see Supplementary Table S3). C_{temp} is allowed to vary between 0.0014 and 0.0030 m °C^{-1} during the calibration, a considerably smaller range than allowed in e.g. HBV-H19 (0.0001 – 0.01 m °C^{-1}). The inclusion of a radiative term makes it possible for the simulated snow cover to respond to changes in incoming shortwave and longwave radiation, as well as surface albedo. Further, adding a radiative term makes it less likely that an unreasonably large amount of snow remains over the summers, ultimately building up so-called 'snow towers' (see e.g. Figure 6 in Skaugen & Weltzien (2016)). The adjustment makes it possible to run the model over consecutive years without zeroing out snow at the beginning of the hydrological year, which has been a common procedure to get rid of 'snow towers' in similar models (Skaugen & Weltzien 2016). The Supplementary Material includes an example of aggregated SWE in a simplified snow module when a traditional degree-day factor, with a constant melt rate of 2.5 mm °C^{-1} day^{-1} is employed, and when the new, radiation-restricted degree-day factor is used (Equation (1)).

The snow module of the HBV model is further modified. The former versions of the model decompose grid cells where ground snow is present into 1–9 sub-grid tiles, depending on the total grid cell SWE, where each tile represents similar snow depths within the grid cell; and further, the numerical snow scheme is applied separately to each of the tiles. In conjunction with adding a radiation-restricted snow-melt routine, the traditional, log-normal SWE-based grid cell tiling is replaced with a simple sigmoidal tanh-function representing grid cell snow cover fraction ($f_s = \tanh(75\,\mathrm{SWE})$, similar to e.g. Roesch *et al.* (2001). Omitting tiling simplifies the model structure, which, following the implementation of the radiation-restricted degree-day factor, would also have needed a tile-area-based discretization of net radiation. Furthermore, the previously used tiling approach was also associated with the build-up of 'snow towers' (Frey & Holzmann 2015).

A final, physically based model enhancement was implemented, which has its base in the availability of high-quality humidity estimates. The traditionally, often calibrated, precipitation phase threshold temperature was replaced with a diagnostic criteria where precipitation is set as snowfall if T2 is below 1 °C and the 2-meter dew point temperature is below 0 °C, and as rain otherwise. Recent studies (Jennings *et al.* 2018; Jennings & Molotch 2019) have shown that including humidity as a predictor of precipitation phase increases its accuracy, as snowfall is more likely in drier rather than more humid environments given the same T2.

Calibration and validation

The model was calibrated with the aim to minimize the regional mean catchment bias in discharge (simulated – observed daily discharge) and to maximize the regional mean KGE. KGE is related to the Nash–Sutcliffe efficiency (NSE), but avoid two caveats embedded within NSE, namely (i) that in order to reach a maximum NSE, variability has to be underestimated, and (ii) within NSE, bias is scaled by the observed temporal standard deviation, which may inflate scores in watersheds with a high seasonal component. The KGE version used is from Gupta *et al.* (2009):

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\sigma_s/\sigma_o - 1)^2 + (\mu_s/\mu_o - 1)^2} \qquad (2)$$

where r is the Pearson correlation coefficient, σ_s and σ_o are the simulated and observed standard deviation, and μ_s and μ_o are the simulated and observed mean. The optimal KGE, and highest possible value, is unity, while it has no lower limit. The KGE is a relative bias measure, so the two optimization goals used in model calibration are not entirely independent.

The model was jointly calibrated for catchments within each of the five calibration regions. The regions following HBV-H19 and are shown in Figure 1(b). Observed daily discharge during the period 2000–2010 from 85 catchments located across mainland Norway was used in the calibration, and discharge observations from 34 independent stations were used for validation. At all calibrated stations, the model was additionally validated for an independent period, 1980 through 1999.

For each of the five regions, only three above-ground parameters were calibrated: a multiplicative correction factor for precipitation, an additional multiplicative

undercatch correction factor in case of snowfall, and the snow-melt degree-day factor (C_{temp}). The six soil parameters described in HBV-B03, FC, beta, perc, kuz, klz, and alpha were individually calibrated for each soil class, except for the glacier and till class, which were merged. Thus, the parameters of a maximum of six soil classes were calibrated for each region, depending on the number of soil classes represented within a region. Lower zone lake runoff response (klz) was set according to that of soil class bog. The regional calibration was conducted using PEST: Model-Independent Parameter Estimation and Uncertainty Analysis (Doherty 2015).

Model evaluation

In order to assess the HBV-E20 model's ability to simulate relevant hydrological states, fluxes and temporal dynamics, results were evaluated against a variety of observational data sets: (i) Simulated SWE was compared with 24,148 SWE observations from 1,181 unique locations from around the time of maximum SWE. (ii) The routine for estimating potential evaporation (E_p), including surface resistance, was evaluated by comparing to May–September mean monthly pan evaporation observations (1967–1972). E_p was calculated for the land cover class 'Open', which represents short vegetation (height = 20 cm, leaf area index = 2, see Supplementary Table S3). The estimated E_p was obtained for the five closest years for which the HBV forcing data is available, i.e. 1979–1984. (iii) The model's ability to reproduce observed trends in monthly discharge was evaluated using a modified Mann–Kendall test including a trend free pre-whitening method (Yue & Wang 2002), as implemented in pyMannKendall (Hussain & Mahmud 2019). For the trend test to be applied 29 out of 30 complete years of daily observations were required, following in-filling of up to two consecutive days of missing data by linear interpolation.

RESULTS

The model was calibrated and validated in terms of regional mean KGE and bias of its daily discharge estimates. The calibrated model parameters are presented in the Supplementary Material. The calibrated parameters were either derived for each calibration region or for each soil class within a calibration region. The calibration provided values for the full study domain except for parameters for the soil class 'marine' within region 1 and 2. These were given calibrated parameters from the soil class 'fluvial'. The model calibration resulted in minor correction factors for rain and snowfall, with the product of the two correction factors amounting to an average increase of 3%. The model corrected precipitation was just 1.4% larger than the SeNorge2018 precipitation between 1980 and 2014 (see Supplementary Material). The model's ability to reproduce potential evaporation, winter maximum SWE, and discharge trends are evaluated below. Finally, the physically enhanced HBV model is applied to provide a mean water balance for the period 1980–2014.

Daily discharge, potential evaporation, maximum SWE, and monthly discharge trends

Daily discharge

The results of the model calibration and validation are given in Table 1, including only catchments where at least 5 years of observations are available. Daily mean and median KGE, with the median given within brackets, mean bias, mean observed discharge, and relative bias are provided. The calibration resulted in a median KGE of 0.74. A KGE above 0.6 was achieved in 86% of the catchments, values above 0.7 in 66% of the catchments, while a KGE score above 0.8 was achieved in 22% of the catchments. In five catchments, the KGE score was above 0.85. The mean bias was −0.3 mm/day or −1.9% of mean discharge. When the model was run for an independent time period, 1980–1999, the median KGE for the catchments was 0.75, while the mean bias was −0.1 mm/day or −0.7% of the mean observed discharge. The model was further evaluated for 34 independent catchments not included in calibration. For these catchments, the median KGE was 0.78, and the mean bias −0.8 mm/day, or −4.8% of observed discharge, for the period 1980–2010. The distribution of KGE and mean bias for the calibration and independent catchments, and for the calibration period and the independent time period is displayed in Supplementary Figure S10. Depending

on the metric considered HBV-E20 show equally good or better performance than previous gridded HBV versions applied for the same domain (see Huang *et al.* 2019 and references therein).

Winter maximum SWE

SWE observations from around the time of maximum SWE are in Figure 3 compared with the model's simulated SWE. The observations and model estimates show a Pearson correlation coefficient of 0.78 (Figure 3(a)). The model's simulated maximum SWE is on average 6 cm higher than the observations; however, as seen in Figure 3(b), there is a strongly significant, positive correlation ($p < 0.000$) between SWE bias and the difference between model grid cell and measurement altitude.

Potential evaporation

The new scheme to calculate potential evaporation, E_p, is compared with pan evaporation in the growing season from May through September in Figure 4. The maps (upper row) and box plots (lower row) compare mean monthly simulated E_p for the land cover class 'Open' (1979–1984) to the pan measurements (1967–1972). A reasonable agreement is found; however, the simulated E_p tends to show higher values than the pan measurements in July, but lower values in September. Overall, the pan measurements show a May to September evaporation of 284 mm, while the simulated values give 272 mm, i.e. 96% of the measured values.

Discharge trends

Trends in observed and simulated monthly discharge for the period 1985–2014 at various measurement stations are shown as a heat map in Figure 5, where a blue color indicates an increasing trend and a red color indicates a decreasing trend. The trends' significance is evaluated using a Mann–Kendall test, and significant trends (low p-values) are marked with one or more asterisks in the plot. At most of the stations, a notable shift in discharge can be seen from early summer or late spring to earlier in spring. This is evident in both the model simulations and observations; the largest monthly change in both is an increase in discharge of around 1.6% in April. The single, largest observed change is an increase of 5.2% at the

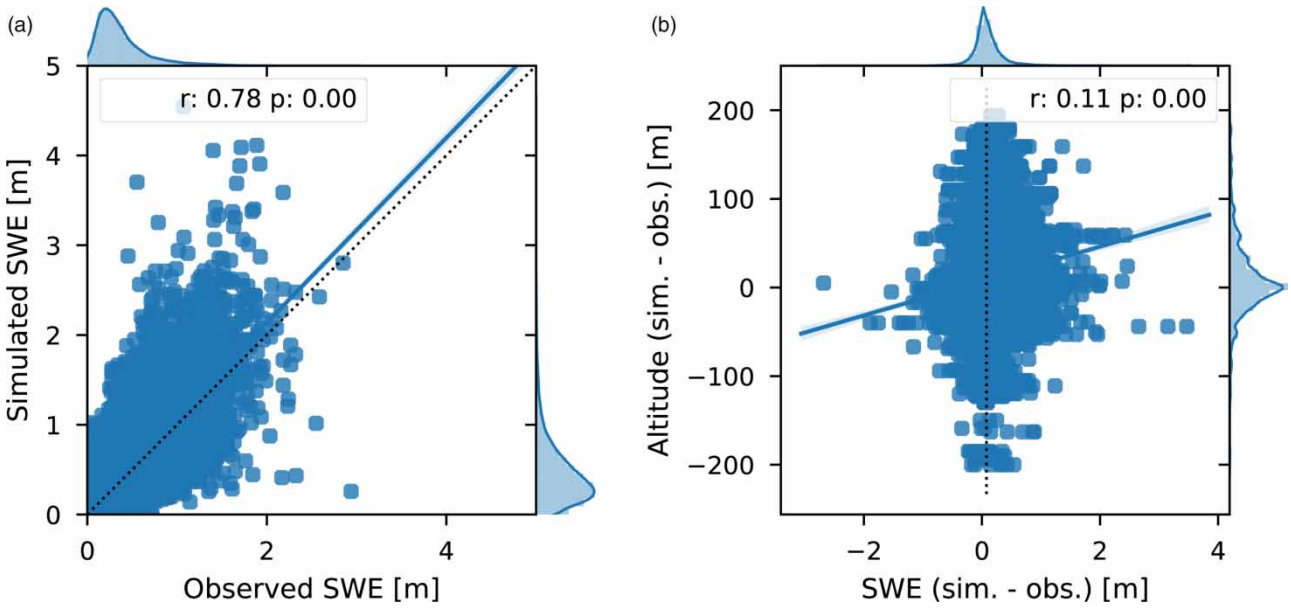


Figure 3 | Observed (x) and simulated (y) SWE around the time of maximum SWE (a), and difference between simulated and observed SWE (x) plotted against difference in altitude between the model grid cell and observation (y) (b). The Pearson correlation and its significance are denoted near the top of each plot.

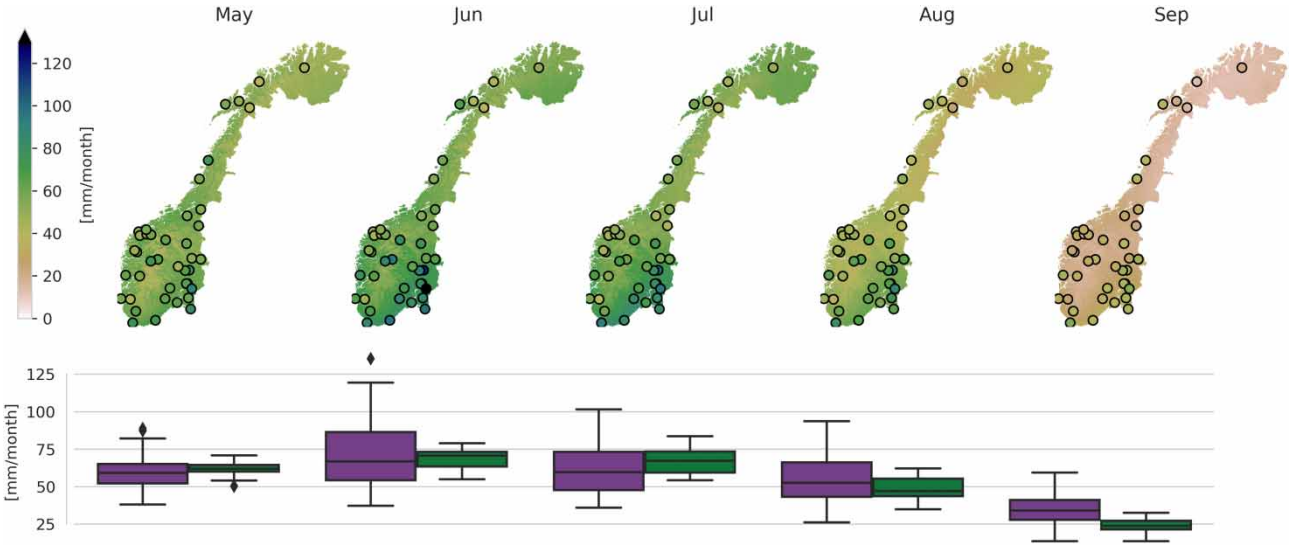


Figure 4 | Mean monthly pan evaporation measured at 42 sites across Norway (1967–1972) shown within black circles overlaid the simulated E_p for the land cover class 'Open' (1979–1984). The lower box plots show the pan measurements in purple and the simulated E_p for collocated grid cells in green. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/nh.2021.132.

station Sundbyfoss, where the model shows a corresponding 3.3% increase. Overall, the trend patterns in the observations are reproduced by the model; however, there are some regions and times of the year with discrepancies, e.g. an observed increase in winter discharge at several stations in Finnmark and Nordland and at Bjoreio (50.13, Hardanger county), which are not reproduced by the model.

1980–2014 water balance

The September 1980 to August 2014 simulated mean annual E_p, evaporation (E), precipitation (P), runoff (R), and the evaporation fraction of precipitation, i.e. E/P, are depicted in Figure 6. There are large regional variations in the water balance elements, with the coastal regions receiving the most precipitation and producing the highest runoff, while more continental regions show the highest potential and actual evaporation. Areal median (mean) annual P is 1,168 (1,367) mm, while R is 975 (1,179) mm, E 157 (178) mm, and E_p 182 (210) mm. The evaporation fraction is 0.17, i.e. just above one-sixth; however, in southeastern and northeastern parts of Norway, annual evaporation reach 40% of precipitation.

DISCUSSION

Model calibration and validation

The herein alterations to the gridded HBV model code implemented in HBV-E20, including a modified physically based potential evaporation routine, a radiation-restricted degree-day factor for snow-melt, and the introduction of humidity as a predictor of precipitation phase led to fewer calibrated parameters. A reduction in the number of parameters reduces the dimensions in parameter space and with that parameter uncertainty, contributing to a more robust model in general, and for climate change impact assessment in particular.

The model achieved a median KGE between 0.74 and 0.78, and a mean bias between −0.1 and −0.8 mm/day, depending on the period and catchments considered. Though the primary goal of the alterations was to increase the physical robustness of the model under climate change, we find that the HBV-E20 model's performance in simulating daily discharge is equally good or better than previous gridded HBV versions applied for the same domain (see Huang *et al.* 2019 and references therein). Additionally, the model showed a fair amount of skill in reproducing observed annual maximum SWE, with a correlation of 0.78 and a

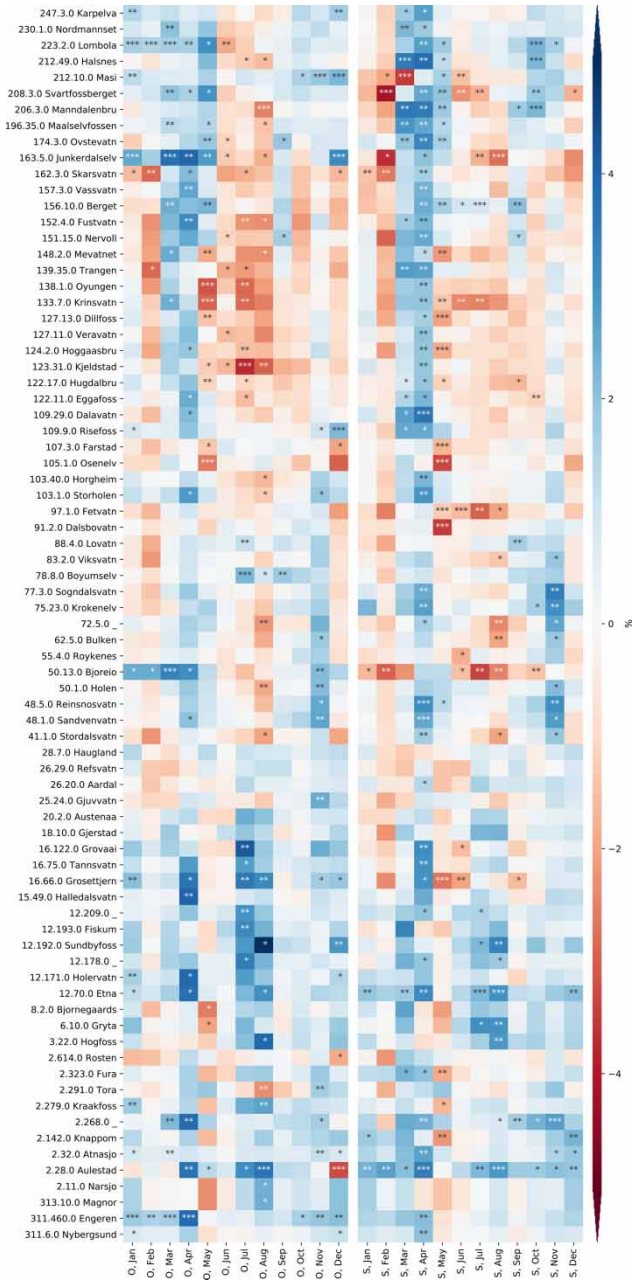


Figure 5 | Observed (left panel) and simulated (right panel) monthly trends in discharge (%) between 1985 and 2014 for individual catchments, sorted from north (upper rows) to south (lower rows). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/nh.2021.132.

mean difference of 6 cm to the point observations. The improved availability of high-quality forcing data facilitated the implementation of model alterations and thus likely contributed to the improved model performance.

A reasonable agreement was seen between pan evaporation measurements (1967–1972) and E_p calculated for the land cover class 'Open' (representing short vegetation) (1979–1984), with the latter amounting to 95% of the former. The pan evaporation measurements show higher regional variability than the estimates. This may, in part, be explained by the fact that the measurements represent points in the terrain with a varying degree of exposure (Hetager & Lystad 1974), while the estimates are provided for 1×1 km grid cell averages. The pan measurements showed relatively higher values in September, whereas the calculated E_p showed relatively higher values in July. Evaporation rates from pans are on average above that of larger, natural water bodies, since pans are surrounded by drier areas, leaving the air less saturated (the oasis effect), and because the pan itself can absorb heat and sunlight, ultimately increasing evaporation. According to Allen *et al.* (1998), reference crop evaporation may be 0.5–1.1 of that measured from a pan, depending on the wind speed, humidity, fetch, and surrounding vegetation. The calculated E_p for the 'Open' land cover class should, similarly to reference crop evaporation, show lower values than the pan measurements. While the variation in the time periods considered may influence the differences seen in observed and simulated E_p, it is possible that the HySN5 incident shortwave radiation, which shows slight overestimations compared with surface observations, contributes to the slightly larger calculated E_p than the pan observations mid-summer. However, the lack of surface observations of incident longwave radiation limits a comprehensive validation of the forcing data estimates, and it is thus difficult to say if the total incident radiation is overestimated or not. A likely larger impact on the calculated E_p is the choice of E_p parameter values for the land cover class 'Open'. The values chosen for the current implementation in HBV should, however, be reasonable, given the close agreement found with the pan measurements. It should be noted that the veracity of estimated E_p for other vegetation types, and the estimates of evaporation from the coniferous forest, where interception loss plays a significant part, remains uncertain due to the lack of observations to constrain the estimates.

A comparison of monthly discharge trends revealed an overall good agreement of the trend; however, for some catchments and calendar months, the simulated and

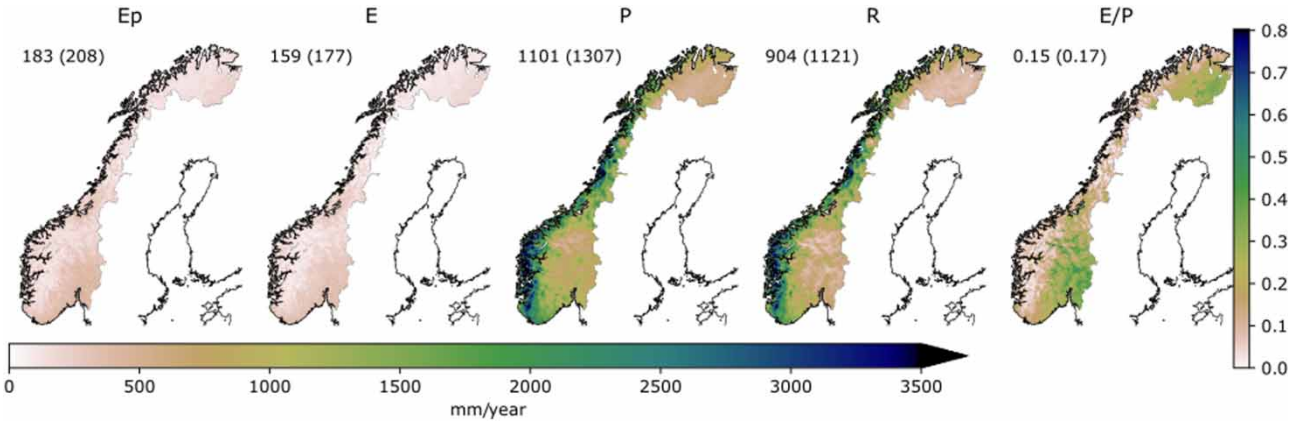


Figure 6 | The 1980–2014 mean annual potential evaporation (E_p), evaporation (E), precipitation (P), runoff (R), and E/P, in mm/year. The areal median and mean are denoted in the upper left corner, with the mean in parenthesis. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/nh.2021.132.

observed trends did not agree. These discrepancies, where present, might be explained by deficiencies in the model, e.g. in its diagnosis of precipitation phase, melting temperature, or perhaps its lack of representation of land use change (e.g. Erlandsen *et al.* 2017). The lack of precipitation observations in some regions, particularly in the mountains, and the variation in the station network feeding into the SeNorge precipitation data set with time, likely limits the ability of the HBV model to reproduce observed discharge trends. Work is currently in progress to resolve the latter issue; a gridded precipitation data set where the observation network is consistent with time (see e.g. Masson & Frei 2016) is under construction (pers. comm. C. Lussana, MET Norway).

The model calibration resulted in very small correction factors for rain and snowfall; model corrected P was just 1.4% larger than the SeNorge2018 precipitation. Allowed ranges of precipitation and snow correction factors under calibration typically vary from around 0.5, i.e. a halving, to 2–3, that is doubling or tripling the precipitation amount (e.g. Table 1 within Huang *et al.* (2019)). The small correction factors of the current study can likely largely be attributed to the high quality of the SeNorge2018 forcing data, which, unlike its predecessor, SeNorge2, includes a correction for precipitation undercatch. The mean annual P for Norway is about 280 mm or 26% higher in SeNorge2018 than its predecessor, SeNorge2. The physically based E_p estimates, being considerably smaller than those estimated in previous versions of the HBV model (see the discussion below), also contribute to smaller precipitation correction factors.

1980–2014 water balance

Averaged over continental Norway, HBV-E20 produced a September 1980 to August 2014 mean annual water balance of 1,367 mm P, E of 178 mm, and 1,179 mm R. Previous water balance estimates for Norway vary considerably (see Table 2). HBV-H19 gives 1983–2012 average annual water balance elements of about 1,333 mm P, 221 mm E, and 1,121 mm R. The climate synthesis report for Norway (Hanssen-Bauer *et al.* 2009, in Norwegian) states that 1961–1990 areal average P is 1,486 mm, E 346 mm, and R 1,140 mm; while the similar, updated climate report published in 2017 (Hanssen-Bauer *et al.* 2017) states that 1971–2000 average P is 1,600 mm, E slightly less than 500 mm, and R 1,100 mm.

Table 2 | Estimates of the annual surface water balance, provided for annual precipitation (P, mm), evaporation (E, mm), and runoff (R, mm), over continental Norway

Source	Period	*P*	*E*	*R*
HBV-E20	1980–2014	1,367	178	1,179
Huang *et al.* (2019) HBV-H19	1983–2012	1,333	221	1,112
Hanssen-Bauer *et al.* (2017) HBV-B03	1971–2000	1,600	500	1,100
Hanssen-Bauer *et al.* (2009) HBV-B03	1961–1990	1,486	346	1,140

It should be noted that the estimates provided in the two climate synthesis reports were based on a simple, temperature-dependent E approximation (as outlined in HBV-B03), and that the estimates are for different reference periods. Furthermore, the precipitation estimates have different sources. The estimates provided in Hanssen-Bauer *et al.* (2009) were based on stations measurements that were interpolated and given correction factors within the gridded hydrological model. The estimates in Hanssen-Bauer *et al.* (2017) were based on a previous version of the national, gridded, observation-based precipitation estimates (SeNorge), and precipitation correction factors added during calibration. The SeNorge data set likely overestimated precipitation (Saloranta 2012). The P estimate of HBV-H19 was based on SeNorge2, after precipitation correction factors had been added.

While the various R estimates are similar, likely since this estimate is constrained by discharge measurements, the estimates of P and E vary considerably. New and improved observational data, particularly of evaporation, incident longwave radiation, and high-altitude precipitation, would help constrain the water balance estimates. Furthermore, future studies regarding the treatment of intercepted precipitation, particularly given that the model is run with a daily resolution, are needed to reduce uncertainty regarding Norway's mean water balance (Tallaksen *et al.* 1996; Haddeland *et al.* 2006). Examples of physically based model enhancements that could help constrain long-term water balance estimates include, e.g.: (i) a sub-daily temporal resolution, which could make way for an improved representation of rain- and snowfall interception; (ii) improved description of sublimation processes; and (iii) inline computation of surface net longwave radiation, possibly based on the inclusion of snow-pack thermal inertia (cold content).

CONCLUSIONS

With the availability of updated and improved forcing data, a gridded version of the HBV model has been enhanced to include a physically based evaporation scheme, a net radiation-restricted degree-day factor approach for snow-melt, and a prescribed precipitation phase threshold based on 2-meter temperature and humidity. The improved forcing data combined with more physically based parameterizations allowed for the model to be calibrated for mainland Norway with fewer free parameters, i.e. only three calibrated above-soil parameters for each of five calibration regions. The model calibration resulted in relatively small correction factors for rain and snowfall, increasing precipitation with just 1.4% compared with the original field. The model showed equally good or better results than previous gridded versions constructed for the same domain. Furthermore, annual maximum SWE, E_p, and discharge trends were satisfactorily represented by the model when compared with observations. The model's mean annual water balance showed lower P and E values for mainland Norway than previous estimates. Additional modeling studies and more observational data are needed to get higher confidence in current and recent estimates of Norway's water balance.

These are the first steps among several that might be undertaken to improve process representation in the model providing more robust long-term water balance estimates in a changing climate. Additional constraints, which might further improve and constrain the physical process description and calibration of the model, include enhanced gridded precipitation data sets, further improved representation and parameterization of the land surface (soil, bedrock, and vegetation), and additional physically based model enhancements.

ACKNOWLEDGEMENTS

The authors would like to thank Tumo Saloranta at NVE for providing SWE data. S.B., S.E., and S.H. were supported by the Norwegian Research Council under grant no. 243803/E10 (I:CAN – Impacts: Climate, Anthroposphere and Nature) and grant no. 295128/E10 (The human imprint on land-atmosphere exchange in high latitudes). The study has further been supported by the Norwegian Centre for Climate Services. This work is a contribution to the Strategic Research Initiative 'Land Atmosphere Interaction in Cold Environments' (LATICE) of the University of Oslo. We also thank two anonymous reviewers for helpful comments on the manuscript.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available. Readers should contact the corresponding author for details.

REFERENCES

Ahlström, A., Bjørkelo, K. & Frydenlund, J. 2014 AR5 klassifikasjonssystem - Klassifikasjon av arealressurser Rapport fra skog og landskap (06/14).

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. 1998 Crop evapotranspiration: Guidelines for computing crop requirements. Irrigation and Drainage Paper No. 56, FAO. https://doi.org/10.1016/j.eja.2010.12.001.

Beldring, S. 2002 Runoff generating processes in boreal forest environments with glacial tills. *Nordic Hydrology*. https://doi.org/10.2166/nh.2002.0013.

Beldring, S., Engeland, K., Roald, L. A., Sælthun, N. R. & Voksø, A. 2003 Estimation of parameters in a distributed precipitation-runoff model for Norway. *Hydrology and Earth System Sciences* **7** (3), 304–316. https://doi.org/10.5194/hess-7-304-2003.

Bruland, O., Sand, K. & Killingtveit, Å. 2001 Snow distribution at a high arctic site at Svalbard. *Nordic Hydrology*. https://doi.org/10.2166/nh.2001.0001.

Buckland, S. M., Grime, J. P., Hodgson, J. G. & Thompson, K. 1997 A comparison of plant responses to the extreme drought of 1995 in Northern England. *The Journal of Ecology*. https://doi.org/10.2307/2960608.

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C., Swenson, S. C. & Zeng, X. 2015 Improving the representation of hydrologic processes in Earth System Models. *Water Resources Research* **51** (8), 5929–5956. https://doi.org/10.1002/2015WR017096.

Cosgrove, B. A. 2003 Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research* **108** (D22), 8842. https://doi.org/10.1029/2002JD003118.

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. & Vitart, F. 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137** (656), 553–597. https://doi.org/10.1002/qj.828.

Doherty, J. 2015 *Calibration and Uncertainty Analysis for Complex Environmental Models*. Watermark Numerical Computing, Brisbane, Australia.

Erlandsen, H. B., Haddeland, I., Tallaksen, L. M. & Kristiansen, J. 2017 The sensitivity of the terrestrial surface energy and water balance estimates in the WRF model to lower surface boundary representations: a South Norway case study. *Journal of Hydrometeorology* **18** (1), 265–284. https://doi.org/10.1175/JHM-D-15-0146.1.

Erlandsen, H. B., Tallaksen, L. M. & Kristiansen, J. 2019 Merits of novel high-resolution estimates and existing long-term estimates of humidity and incident radiation in a complex domain. *Earth System Science Data* **11** (2), 797–821. https://doi.org/10.5194/essd-11-797-2019.

Ferguson, R. I. 1999 Snowmelt runoff models. *Progress in Physical Geography: Earth and Environment* **23** (2), 205–227. https://doi.org/10.1177/030913339902300203.

Frey, S. & Holzmann, H. 2015 A conceptual, distributed snow redistribution model. *Hydrology and Earth System Sciences* **19** (11), 4517–4530. https://doi.org/10.5194/hess-19-4517-2015.

Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. 2009 Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *Journal of Hydrology* **377** (1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Haddeland, I., Lettenmaier, D. P. & Skaugen, T. 2006 Reconciling simulated moisture fluxes resulting from alternate hydrologic model time steps and energy budget closure assumptions. *Journal of Hydrometeorology* **7** (3), 355–370. https://doi.org/10.1175/JHM496.1.

Hanssen-Bauer, I., Drange, H., Førland, E., Roald, L. A., Børsheim, K. Y., Hisdal, H., Lawrence, D., Nesje, A., Sandven, S., Sorteberg, A., Sndby, S., Vasskog, K. & Ådlandsvik, B. 2009 *Klima i Norge 2100 Bakgrunnsmateriale til NOU Klimatilpasning (Climate in Norway 2100 Background Material for NOU Climate Adaptation)*, p. 148.

Hanssen-Bauer, I., Førland, E. J., Haddeland, I., Hisdal, H., Lawrence, D., Mayer, S., Nesje, A., Nilsen, J. E. Ø., Sandven, S., Sandø, A. B., Sorteberg, A. & Ådlandsvik, B. 2017 Climate in Norway 2100 – a knowledge base for climate adaptation. Norwegian Environment Agency (Miljødirektoratet).

Hegdahl, T. J., Tallaksen, L. M., Engeland, K., Burkhart, J. F. & Xu, C. Y. 2016 Discharge sensitivity to snowmelt parameterization: a case study for Upper Beas basin in Himachal Pradesh, India. *Hydrology Research*. https://doi.org/10.2166/nh.2016.047.

Hersbach, H., Rosnay, P. d., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Alonso-Balmaseda, M., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J.-R., de Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D., Dragani, R. & Zuo, H. 2018 *Operational Global Reanalysis: Progress, Future Directions and Synergies with NWP*. ERA Report Series (ERA Report).

Hetager, S. E. & Lystad, S. L. 1974 *Fordampning fra fri vannflate: verdier basert på målinger i perioden 1967–1972*. Den Norske komité for Den internasjonale hydrologiske dekade.

Hrachowitz, M. & Clark, M. P. 2017 HESS opinions: the complementary merits of competing modelling philosophies

in hydrology. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-21-3953-2017

Huang, S., Eisner, S., Magnusson, J. O., Lussana, C., Yang, X. & Beldring, S. 2019 Improvements of the spatially distributed hydrological modelling using the HBV model at 1 km resolution for Norway. *Journal of Hydrology* **577**, 123585. https://doi.org/10.1016/J.JHYDROL.2019.03.051.

Hussain, M. & Mahmud, I. 2019 pyMannKendall: a python package for non parametric Mann Kendall family of trend tests. *Journal of Open Source Software* **4** (39), 1556. https://doi.org/10.21105/joss.01556.

Jennings, K. S. & Molotch, N. P. 2019 The sensitivity of modeled snow accumulation and melt to precipitation phase methods across a climatic gradient. *Hydrology and Earth System Sciences* **23** (9), 3765–3786. https://doi.org/10.5194/hess-23-3765-2019.

Jennings, K. S., Winchell, T. S., Livneh, B. & Molotch, N. P. 2018 Spatial variation of the rain-snow temperature threshold across the Northern Hemisphere. *Nature Communications* **9** (1), 1148. https://doi.org/10.1038/s41467-018-03629-7.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P. & Thielen, J. 2016 Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling and Software*. https://doi.org/10.1016/j.envsoft.2015.09.009.

Kelliher, F. M., Leuning, R., Raupach, M. R. & Schulze, E. D. 1995 Maximum conductances for evaporation from global vegetation types. *Agricultural and Forest Meteorology*. https://doi.org/10.1016/0168-1923(94)02178-M.

Körner, C. 1995 Leaf diffusive conductances in the major vegetation types of the globe. *Ecophysiology of Photosynthesis*. https://doi.org/10.1007/978-3-642-79354-7_22.

Kustas, W. P., Rango, A. & Uijlenhoet, R. 1994 A simple energy budget algorithm for the snowmelt runoff model. *Water Resources Research*. https://doi.org/10.1029/94WR00152.

Leuning, R., Zhang, Y. Q., Rajaud, A., Cleugh, H. & Tu, K. 2008 A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation. *Water Resources Research*. https://doi.org/10.1029/2007WR006562.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. 1997 Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* **201** (1–4), 272–288. https://doi.org/10.1016/S0022-1694(97)00041-3.

Lussana, C., Tveito, O. E. & Uboldi, F. 2018 Three-dimensional spatial interpolation of 2 m temperature over Norway. *Quarterly Journal of the Royal Meteorological Society* **144** (711), 344–364. https://doi.org/10.1002/qj.3208.

Lussana, C., Tveito, O. E., Dobler, A. & Tunheim, K. 2019 seNorge_2018, daily precipitation, and temperature datasets over Norway. *Earth System Science Data* **11** (4), 1531–1551. https://doi.org/10.5194/essd-11-1531-2019.

Magnusson, J., Wever, N., Essery, R., Helbig, N., Winstral, A. & Jonas, T. 2015 Evaluating snow models with varying process representations for hydrological applications. *Water Resources Research*. https://doi.org/10.1002/2014WR016498.

Majasalmi, T., Eisner, S., Astrup, R., Fridman, J. & Bright, R. M. 2018 An enhanced forest classification scheme for modeling vegetation–climate interactions based on national forest inventory data. *Biogeosciences* **15** (2), 399–412. https://doi.org/10.5194/bg-15-399-2018.

Masson, D. & Frei, C. 2016 Long-term variations and trends of mesoscale precipitation in the Alps: recalculation and update for 1901–2008. *International Journal of Climatology* **36** (1), 492–500. https://doi.org/10.1002/joc.4343.

Merz, R., Parajka, J. & Blöschl, G. 2011 Time stability of catchment model parameters: implications for climate impact analyses. *Water Resources Research* **47** (2). https://doi.org/10.1029/2010WR009505.

Milly, P. C. D. & Dunne, K. A. 2011 On the hydrologic adjustment of climate-model projections: the potential pitfall of potential evapotranspiration. *Earth Interactions* **15** (1), 1–14. https://doi.org/10.1175/2010EI363.1.

Monteith, J. L. 1965 Evaporation and environment. In *Symposia of the Society for Experimental Biology*, Vol. 19. Cambridge University Press (CUP), Cambridge, pp. 205–234.

Mu, Q., Zhao, M. & Running, S. W. 2011 Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment* **115** (8), 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models: part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Roesch, A., Wild, M., Gilgen, H. & Ohmura, A. 2001 A new snow cover fraction parameterization for the ECHAM4 GCM. *Climate Dynamics* **17** (12), 933–946. https://doi.org/10.1007/s003820100153.

Saloranta, T. M. 2012 Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model. *The Cryosphere* **6** (6), 1323–1337. https://doi.org/10.5194/tc-6-1323-2012.

Schulze, E. D., Kelliher, F. M., Korner, C., Lloyd, J. & Leuning, R. 1994 Relationships among maximum stomatal conductance, ecosystem surface conductance, carbon assimilation rate, and plant nitrogen nutrition: a global ecology scaling exercise. *Annual Review of Ecology and Systematics*. https://doi.org/10.1146/annurev.es.25.110194.003213.

Sheffield, J., Goteti, G. & Wood, E. F. 2006 Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate* **19** (13), 3088–3111. https://doi.org/10.1175/JCLI3790.1.

Skaugen, T. & Weltzien, I. H. 2016 A model for the spatial distribution of snow water equivalent parameterized from the spatial variability of precipitation. *Cryosphere* **10** (5), 1947–1963. https://doi.org/10.5194/tc-10-1947-2016.

Tallaksen, L. M., Schunselaar, S. & Van Veen, R. 1996 Comparative model estimates of interception loss in a

coniferous forest stand. *Nordic Hydrology* **27** (3), 143–160. https://doi.org/10.2166/nh.1996.0001.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J. & Viterbo, P. 2014 The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-Interim reanalysis data. *Water Resources Research* **50** (9), 7505–7514. https://doi.org/10.1002/2014WR015638.

Weynants, M., Montanarella, L. & Tóth, G. 2013 European HYdropedological Data Inventory (EU-HYDI). In: *European Commission Joint Research Centre Institute for Environment and Sustainability*. https://doi.org/10.2788/5936.

Wolff, M. A., Isaksen, K., Petersen-ØVerleir, A., Ødemark, K., Reitan, T. & Brækkan, R. 2015 Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study. *Hydrology and Earth System Sciences* **19** (2), 951–967. https://doi.org/10.5194/hess-19-951-2015.

Wong, W. K., Beldring, S., Engen-Skaugen, T., Haddeland, I. & Hisdal, H. 2011 Climate change effects on spatiotemporal patterns of hydroclimatological summer droughts in Norway. *Journal of Hydrometeorology*. https://doi.org/10.1175/2011JHM1357.1.

Yue, S. & Wang, C. Y. 2002 Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resources Research*. https://doi.org/10.1029/2001wr000861.

Zappa, M., Pos, F., Strasser, U., Warmerdam, P. & Gurtz, J. 2003 Seasonal water balance of an alpine catchment as evaluated by different methods for spatially distributed snowmelt modelling. *Nordic Hydrology*. https://doi.org/10.2166/nh.2003.0003.

First received 5 September 2020; accepted in revised form 11 December 2020. Available online 19 January 2021