Journal of Hydrology 577 (2019) 123585

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Improvements of the spatially distributed hydrological modelling using the HBV model at 1 km resolution for Norway

Shaochun Huang[a,*], Stephanie Eisner[b], Jan Olof Magnusson[a], Cristian Lussana[c], Xue Yang[d], Stein Beldring[a]

[a] Norwegian Water Resources and Energy Directorate (NVE), P.O. Box. 5091 Majorstua, 0301 Oslo, Norway
[b] Norwegian Institute of Bioeconomy Research (NIBIO), P.O. Box. 115, 1431 Ås, Norway
[c] Norwegian Meteorological Institute, Henrik Mohns Plass 1, 0313 Oslo, Norway
[d] Department of Geosciences, University of Oslo, P.O. Box 1047 Blindern, Oslo N-0316, Norway

ARTICLE INFO

This manuscript was handled by Marco borga, Editor-in-Chief, with the assistance of Eylon Shamir, Associate Editor

Keywords:
Regional calibration
Norway
HBV
High resolution
Potential evapotranspiration

ABSTRACT

A robust hydrological modeling at a fine spatial resolution is a vital tool for Norway to simulate river discharges and hydrological components for climate adaptation strategies. However, it requires improvements of modelling methods, detailed observational data as input and expensive computational resources. This work aims to set up a distributed version of the HBV model with a physically based evapotranspiration scheme at 1 km resolution for mainland Norway and to calibrate/validate the model for 124 catchments using regionalized parameterizations. The Penman-Monteith equation was implemented in the HBV model and vegetation characteristics were derived from the Norwegian forest inventory combined with multi-source remote sensing data at 16 m spatial resolution. The estimated potential evapotranspiration (Ep) was compared with pan measurements and estimates from the MODerate Resolution Imaging Spectrometer (MOD16) products, the Global Land Evaporation Amsterdam Model (GLEAM) and Variable Infiltration Capacity (VIC) hydrological model. There are 5 climatic zones in Norway classified based on 4 temperature and precipitation indices. For each zone, the model was calibrated separately by optimizing a multi-objective function including the Nash-Sutcliff efficiency (NSE) and biases of selected catchments. In total, there are 85 catchments for calibration and 39 for validation. The Ep estimates showed good agreement with the measurements, GLEAM and VIC outputs. However, the MOD16 product significantly overestimates Ep compared to the other products. The discharge was well reproduced with the median daily NSE of 0.68/0.67, bias of −3%/−1%, Kling-Gupta efficiency (KGE) of 0.70/0.69 and monthly NSE of 0.80/0.78 in the calibration/validation periods. Our results showed a significant improvement compared to the previous HBV application for all catchments, with an increase of 0.08–0.16 in the median values of the daily NSE, KGE and monthly NSE. Both the temporal and spatial transferability of model parameterizations were also enhanced compared to the previous application.

1. Introduction

Understanding the past, present and future climate impacts serves as basis to develop adaptation strategies for decision makers. It requires high spatial resolution information on projected impacts, which can be directly applied in adaptation measures (Kaspersen et al., 2012). For the water sector, hydrological models are important tools to project the impact of climate change on runoff, water resources and flooding at both regional and large scales (Olsson et al., 2016).

In Norway, the HBV model is one of the most widely used tools for runoff simulations (Bergström, 2006). The spatially distributed version of this model was used to produce the map of average annual runoff for Norway at 1 km resolution (Beldring et al., 2003) and showed good performance for river discharge and hydrological components (e.g. snow storage, soil moisture, groundwater recharge and runoff). In addition, the model has been used to provide fine resolution hydrological projections in the climate impact assessments for the governmental report “Climate in Norway 2100” (Hanssen-Bauer et al., 2017) published by the Norwegian Centre for Climate Services (NCCS).

Even though the HBV models has been used widely and shown good performance during evaluations, the model requires further improvements on its conceptual model structure to better quantify the impacts of both climate and land use changes. There is an increasing discussion whether conceptual rainfall-runoff models are capable to simulate the

* Corresponding author.
E-mail address: shh@nve.no (S. Huang).

https://doi.org/10.1016/j.jhydrol.2019.03.051
Received 31 August 2018; Received in revised form 6 March 2019; Accepted 7 March 2019
Available online 18 March 2019

water balance under changing climatic conditions (Coron et al., 2014; Fowler et al., 2016; Merz et al., 2011). The calibrated parameters may not be valid under non-stationary conditions, and insufficient model calibration and validation strategies may further contribute to poor temporal transferability of models. In addition to such issues, measured precipitation at gauge stations exhibit problems in Norway such as undercatch (Førland et al., 1996). Such uncertainties largely influences estimates of, for example, evapotranspiration (E) by hydrological models. Since E is constrained by potential evapotranspiration (Ep) and soil moisture in the model, robust Ep estimates can help to reduce the uncertainty caused by biased precipitation data. Several studies reveal that different Ep formulas could lead to different changes in Ep as well as river discharge under climate scenarios (McAfee, 2013; Seiller and Anctil, 2016). The use of data-intensive methods, i.e. the methods consider other climatic variables in addition to air temperature, are recommended when estimating the possible effects of climatic changes on evaporative demand (Donohue et al., 2010; McAfee, 2013).

A major restriction of using physically based Ep methods (e.g. Penman-Monteith) on a large scale is the lack of spatially distributed input data, foremost vegetation parameters and meteorological forcing data. Recently, an enhanced forest classification scheme based on the Norwegian forest inventory combined with multi-source remote sensing data (Majasalmi et al., 2018) has become available, providing detailed information at 16 m spatial resolution on forest vegetation characteristics in Norway. This enabled us to estimate Ep using the Penman-Monteith approach in the HBV model and to compare it with pan evaporation measurements and satellite-based products, such as products from the MODerate Resolution Imaging Spectrometer (MOD16) (Mu et al., 2011) and the Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011). Satellite-based products are considered as important land evaporation estimates on terrestrial water and energy cycles and environmental changes (Miralles et al., 2016), and are increasingly used for model calibration or as input to hydrological models (Bowman et al., 2017; Demirel et al., 2018; Lopez et al., 2017; Spies et al., 2015). However, these satellite-derived data have not yet been validated in Norway due to lack of, for example, long-term flux net data. Hence, a comparison between the ground based measurements and the Ep estimates using different input data and methods can improve the understanding of evaporation on the national scale, which has rarely been documented in literature.

Another potential improvement for the distributed version of the HBV modelling can be the application of a new calibration approach. The previous version of the HBV model was calibrated simultaneously against the discharge of 141 catchments across Norway (Beldring et al., 2003). The calibration lasted for three months due to a large number of grid cells (more than 300,000) and 69 calibration parameters. In the recent decades, there is an increasing trend of applying regionalization calibration approaches in large-scale hydrological modelling (Abbaspour et al., 2015; Beck et al., 2016; Hundecha et al., 2016; Troy et al., 2008). One widely used approach is to construct a regression model linking model parameters to catchment characteristics or other variables and perform the calibration (Hundecha et al., 2016; Rakovec et al., 2016). Other common approaches focus on geographic proximity (Oudin et al., 2008), the climatic and/or physiographic similarity of catchments or grid cells (Beck et al., 2016; Parajka et al., 2007; Troy et al., 2008). They either interpolated independently calibrated parameters for selected catchments to other non-gauged areas or calibrated multiple catchments with similar characteristics simultaneously and transfer the parameters to ungauged areas which have the similar characteristics of the calibrated catchments.

Yang et al. (2018) compared the performance of different regionalization approaches using the Water and Snow balance MODeling system (WASMOD) for 118 Norwegian catchments. They found that applying a set of regionalized parameters to catchments within one climatic zone shows similar results as the distance-based regionalization methods and outperforms the regression-based ones. This study motivated us to apply a regionalized approach based on climatic similarities to calibrate the distributed HBV model. It is expected to improve the runoff simulations both temporally and spatially for the whole land surface of Norway. In addition, it can speed up the calibration procedure by calibrating different groups of catchments in parallel.

Thus, this study attempts to improve both the HBV model structure and the calibration procedure to better simulate Ep as well as discharges for mainland Norway. It shows the first application of the modified HBV model using the Penman-Monteith method with vegetation characteristics derived from a high-resolution dataset and a regionalized calibration procedure. The main objectives of this study are formulated as 1) to set up the distributed version of the HBV model with the Penman-Monteith method at 1 km resolution for the mainland of Norway; 2) to evaluate the Ep estimates using both ground and satellite-based data; 3) to calibrate and validate the model for 124 catchments using regionalized parameterizations and 4) to compare the simulated river discharge and hydrological components with the outputs from the previous HBV application used in the “Climate in Norway 2100” report (Hanssen-Bauer et al., 2017).

2. Study area

Norway is located in Northern Europe and covers an area of about 325 000 km^2. The elevation spans from 0 m along the coast to more than 2000 m in the central regions (Fig. 1). About half of the land area is covered by bare soil and heathers in high mountains and about 38% of the area is covered by forest. There is a large number of lakes and bogs/wetlands in Norway, covering about 5% and 6% of the total land surface, respectively. Only about 3% of the land surface is used as arable land. Norwegian soils and aquifers are generally shallow and have a limited storage capacity consisting mainly of thin precipitation-fed till deposits. The bare mountain and moraine soils account for ca. 80% of the total mainland area.

The mean annual temperature during 1971–2000 was about 1.3 °C for the mainland (Hanssen-Bauer et al., 2017). It is warmest along the coast of southern Norway with the annual temperature up to 7 °C while it is coldest in the high mountains with the annual temperature down to −4 °C. Due to the low temperature, only about 11% of the land area has a growing season lasting more than six months (Hanssen-Bauer et al., 2017). Similar to temperature, the annual precipitation also shows a distinct spatial distribution, ranging from 300 mm in south-eastern and northern Norway to more than 3500 mm in the west. The mainland can be divided into 6 hydrological regimes: Finnmark, Nordland, Trøndelag, Western Norway, Eastern Norway and Southern Norway (Fig. 1).

In total, 124 catchments were selected for model calibration and validation in this study (Figs. 1 and 2). The distribution of the catchments represents various climate and hydrological regimes, geographic conditions and landscape types in Norway. These catchments are also used for calibration in the previous HBV application and most of them are considered as near-natural catchments (Beldring et al., 2003). The size of the catchments varies from 7 to 15450 km^2 and 85% of them are smaller than 1000 km^2. As some of the catchments stretch along the national borders, the catchment area in the neighbouring countries was also included in the HBV modelling (see Fig. 2). Detailed information of the 124 catchments is listed in Appendix A.

3. Methods and data

3.1. HBV model

The spatially distributed version of the HBV model used here (Beldring et al., 2003) was developed for calculating the water balance for 1 × 1 km^2 grid cells covering the entire mainland surface area of Norway. Each grid cell can be subdivided into lake area, glacier area

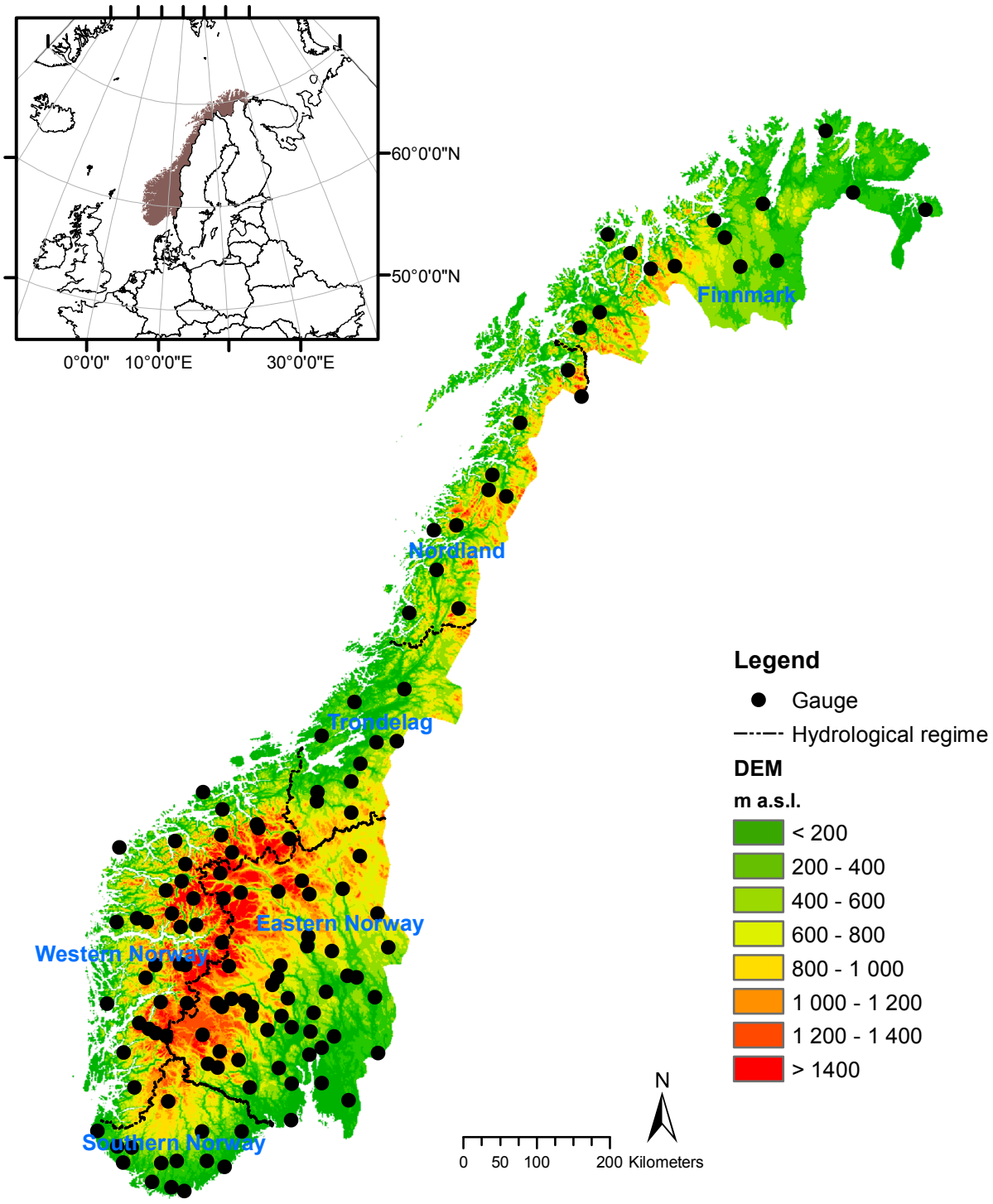


Fig. 1. The Digital Elevation Model (DEM) of the mainland Norway, the location of the 124 discharge gauging stations and the hydrological regimes in Norway.

and up to three vegetation types. The model runs with a daily time step using precipitation and air mean temperature as input. It calculates subgrid scale accumulation and ablation of snow, interception storage, distribution of soil moisture storage, E, groundwater storage, runoff response, lake evaporation and glacier mass balance. The original method for computing Ep uses a function of air temperature and the E is calculated based on Ep, field capacity and permanent wilting point (see Beldring et al., 2003 for details and Table 1 for the range of field capacity). In this study, the permanent wilting point was assumed as zero meter based on the assessments of the HBV parameterizations for Scandinavian catchments by Bergström (1976). The snowmelt is calculated by a simple degree-day method. The runoff is simulated by two non-linear parallel reservoirs representing direct discharge and the groundwater response. The calibration parameters include parameters for evaporation and snowmelt processes in each land use type and for the soil and groundwater processes in each soil type (Table 1). Detailed information of the model and the algorithms can be found in Beldring et al. (2003) and Bergström (1995). We reference the distributed HBV model described by Beldring et al. (2003) as the old version of the HBV model in this study.

In this study, we improve the model structure by replacing the simple Ep function with the Penman-Monteith method. Ershadi et al.

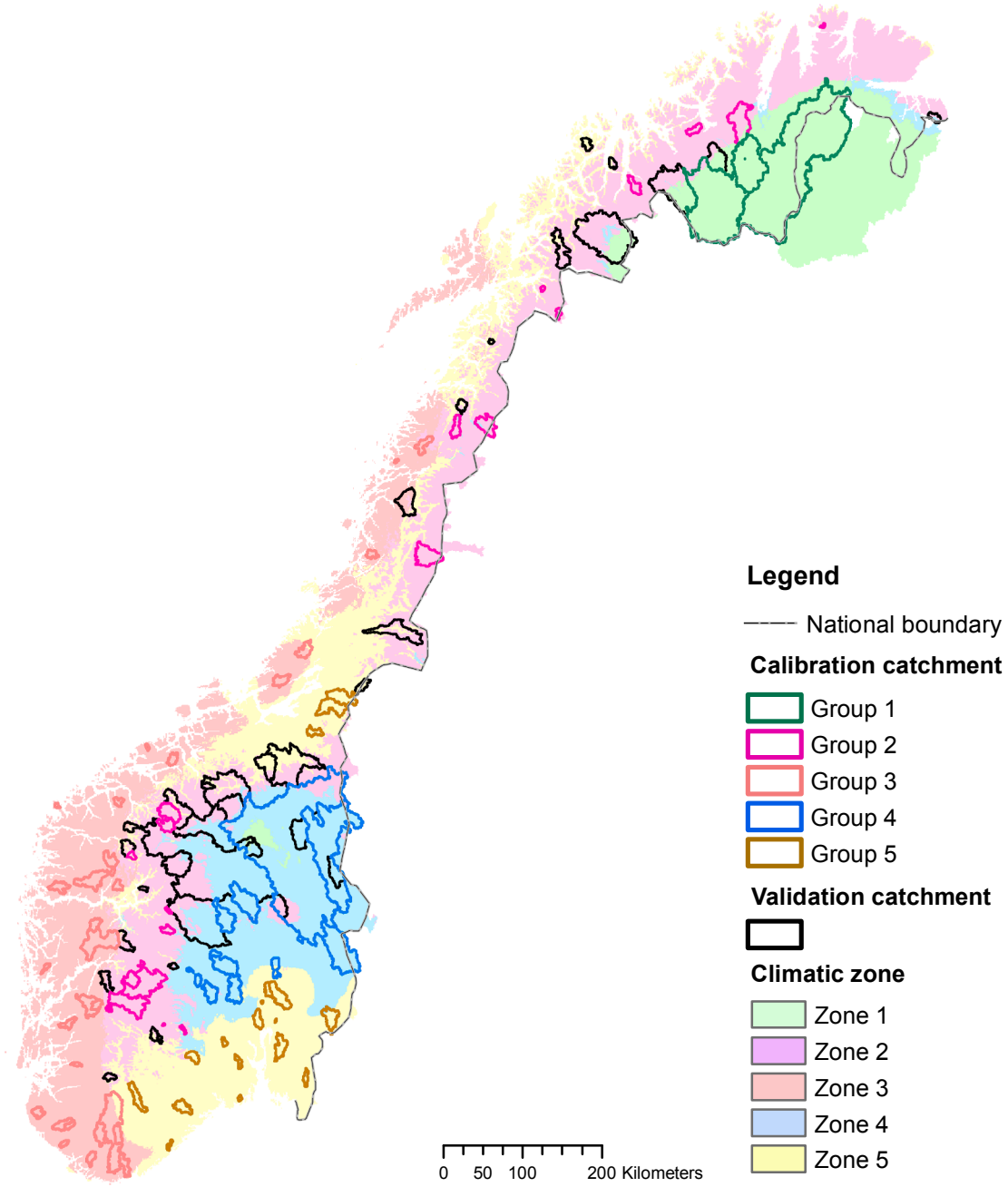


Fig. 2. Map of the modelling domain with the background colours indicating the five different climatic zones. The catchments used for calibration are coloured according to which climatic zone they belong to, whereas watersheds with black borders were used for validation.

(2015) compared the estimates of E using different structures of the Penman–Monteith model, a traditional single-source model (Monteith, 1965), a two-layer model based on Shuttleworth and Wallace (1985) and a three-source model based on Mu et al. (2011). They found that no single model structure consistently outperformed the others for all biome types. In addition, best estimates are often generated using the surface resistance parameterization based on the lookup-table from Mu et al. (2011) and the Thom (1975) formulation for aerodynamic resistance that incorporated dynamic values of roughness parameters.

Based on their findings and data availability of this study, we applied the traditional single-source Penman-Monteith method (Eq. (1)) to estimate *Ep* (mm).

$$\lambda_v E_p = \frac{\Delta(R_n - G) + \rho_a c_p (v_s - v_a)/r_a}{\Delta + \gamma * \left(1 + \frac{r_s}{r_a}\right)} \quad (1)$$

where λ_v is the latent heat of vaporization (J kg^{-1}), R_n is the net radiation (W m^{-2}), G is the soil heat flux (W m^{-2}), $(v_s - v_a)$ represents the vapor pressure deficit of the air (Pa), ρ_a is the density of air at constant pressure (kg m^{-3}), c_p is the specific heat of the air (J kg^{-1} K^{-1}), Δ represents the slope of the saturation vapor pressure temperature relationship (Pa K^{-1}), γ is the psychrometric constant, and r_s and r_a denote the surface and aerodynamic resistances (s m^{-1}).

The Thom formulation was used to estimate r_a (Eq. (2)).

$$r_a = \frac{\ln\left[\frac{z_w - d}{z_{ow}}\right]\ln\left[\frac{z_h - d}{z_{oh}}\right]}{k^2 u_z} \quad (2)$$

where z_w is the height of the wind instrument (m), z_{ow} is the roughness length governing momentum transfer (m), z_h is the height of the humidity measurements (m), z_{oh} is the roughness length governing

Table 1
List of calibration parameters.

	Parameter	Explanation	Unit	Min	Max	Use in old HBV	Used in new HBV
General	CORR_RAIN	Precipitation correction for undercatch	–	0.5	1.5	no	yes
	CORR_SNOW	Snow correction for undercatch	–	0.5	2	no	yes
Lake	KLAKE	Rating curve constant	–	1.00E−04	0.1	yes	yes
	EPOT_PAR_LAKE	Temperature index for potential evaporation from lakes	m/°C	1.00E−06	0.002	yes	no
Land use	INTER_MAX	Maximum interception storage	m	1.00E−04	0.03	yes	no
	EOT_PAR	Temperature index for potential evaporation	m/°C	1.00E−06	0.002	yes	no
	SMELT_T	Snow melt temperature	°C	−1	2	yes	yes
	SMELTR	Temperature index for snow melt rate	m/°C	1.00E−04	0.01	yes	yes
	IMELTR	Additional coefficient to snow melt rate	–	1	4	yes	yes
Soil	FC	Field capacity	m	1.00E−02	1	yes	yes
	BETA	Shape coefficient of soil moisture	–	1	5	yes	yes
	KUZ	Upper zone recession coefficient	–	1.00E−03	1	yes	yes
	ALFA	Upper zone nonlinear drainage coefficient	–	1	2	yes	yes
	PERC	Percolation from upper zone to lower zone	–	1.00E−03	0.5	yes	yes
	KLZ	Lower recession coefficient	–	1.00E−03	1	yes	yes

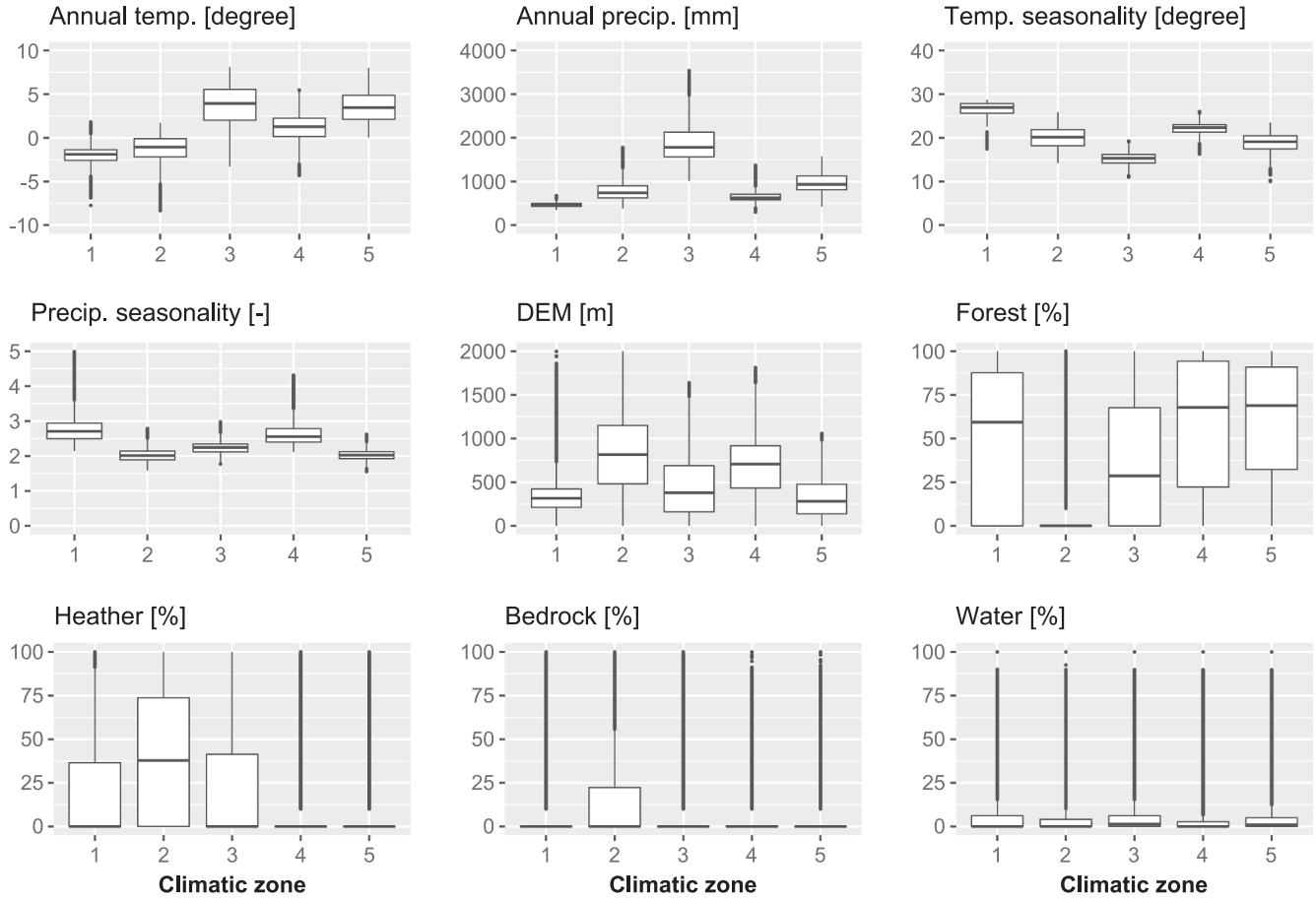


Fig. 3. Climatic, geophysical and land cover characteristics for the grids in the five climatic zones.

transfer of heat and vapour (m) ($z_{oh} = 0.1 z_{om}$), d is the zero plane displacement height (m), k is the Kármán's constant (0.41), u_z is the wind speed at height z_w (m/s) and h is the vegetation height (m). Both the z_{ow} and d are a function of vegetation height and leaf area index (LAI) following Federer et al. (1996).

The LAI varies for deciduous forest and crops seasonally and it is assumed to be constant for other vegetation types. Budburst of deciduous forest is determined by (spring) temperatures using a Growing Degree Day (GDD) sum approach (Olsson and Jonsson, 2014) while the timing of autumn phenophases in a given location is assumed to be determined by photoperiod and it is simulated with constant senescence dates. The LAI for crops is estimated by a simplified EPIC crop model (Williams et al., 1984).

We applied the Biome-Property-Look-Up-Table from Mu et al. (2011) to estimate the surface resistance parameters for vegetation. For other land cover types, e.g. urban and built-up, we used the parameterizations from the evapotranspiration model AMOR (Tallaksen et al., 1996), which was developed for Norway.

All parameterization for the Penman-Monteith method is prescribed in the HBV model (see Appendix B for the prescribed parameter values). Since there is no calibration parameters needed for the Ep simulation, e.g. LAI, vegetation height, albedo and bulk resistance, the Ep estimates

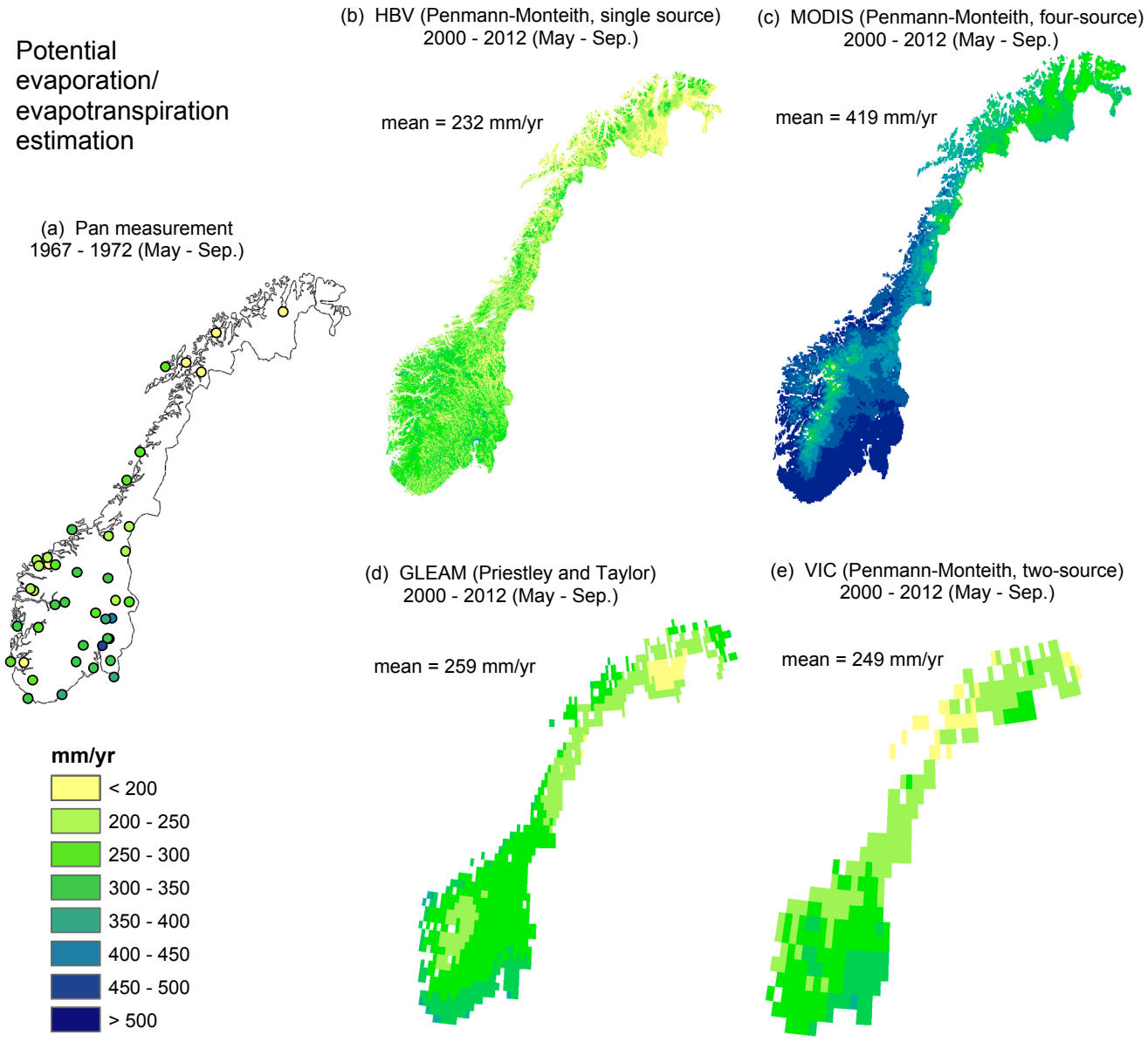


Fig. 4. The average annual potential evaporation measured by the Thorsrud 2500 evaporimeter at 42 stations (a) and average annual potential evapotranspiration for the mainland of Norway estimated by the HBV (b), MOD16 (c), GLEAM (d) and VIC (e) models.

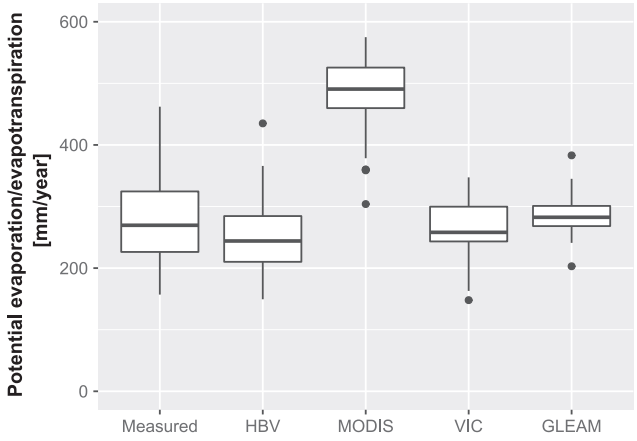


Fig. 5. Measured and simulated annual Ep at the 42 stations equipped with the Thorsrud 2500 evaporimeter.

must be validated against ground measurements and other model products before the model is calibrated against discharges. In addition, we corrected the rainfall gauge undercatch by a continuous adjustment function derived by Wolff et al. (2015) based on data from a test site in southern Norway. However, we found that this correction function cannot guarantee a reliable correction for other parts of Norway, thus the precipitation/snow correction coefficients need to be adjusted for each climatic zone in this study (Table 1).

3.2. Regionalized calibration

The HBV model was calibrated and validated using discharge observations from 124 catchments spread across the whole country (Fig. 2). As we mentioned in Section 3.1, all parameters used by the Penman-Monteith method were prescribed, so we should validate the Ep estimates to ensure a reasonable simulation for this process. Based on the validation of Ep for the period 2000–2012 (the details on the selection of period are explained in Section 3.4), we can start the model calibration against discharges for the period of 2000–2007, which lies

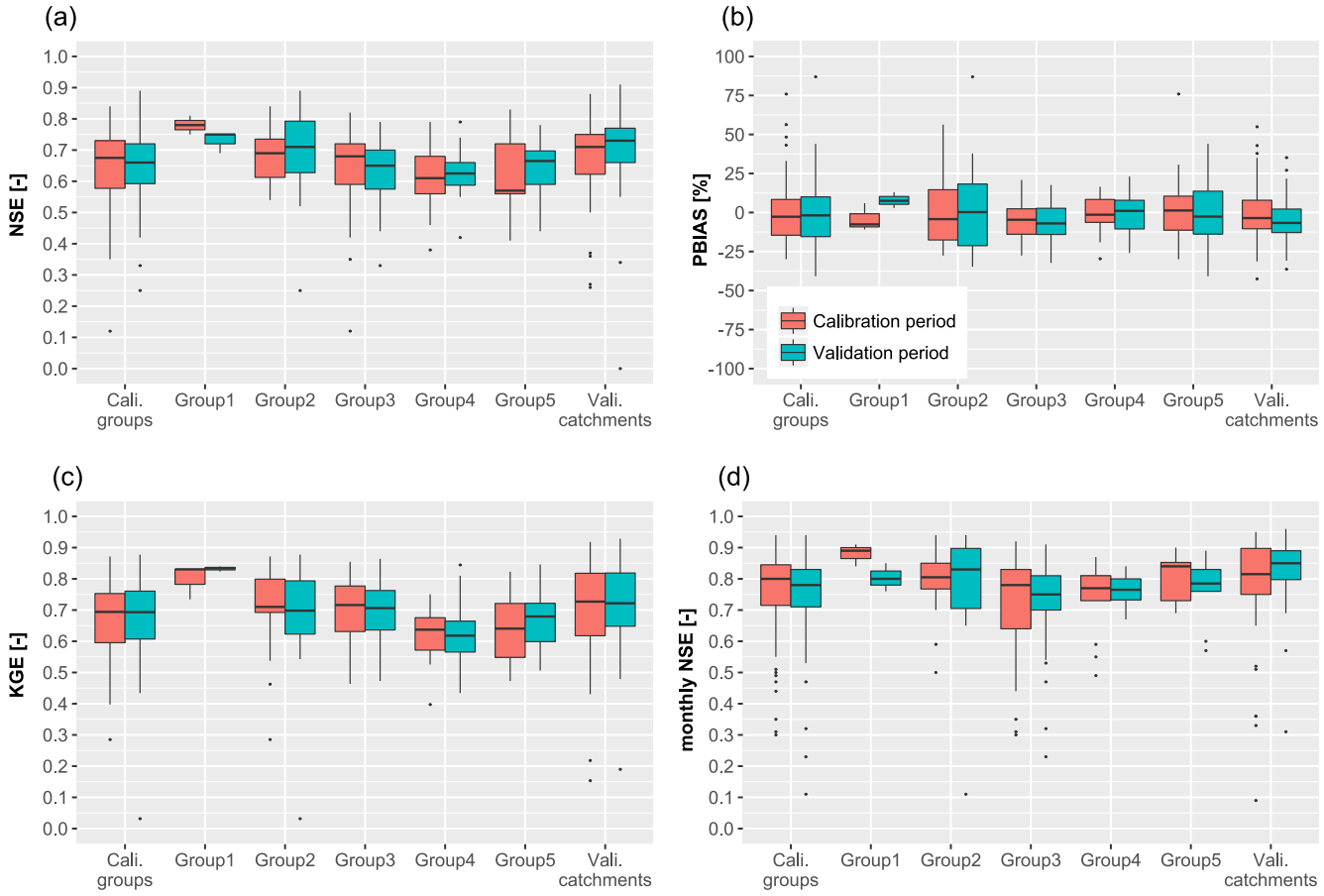


Fig. 6. Model performance for the catchments in all calibration groups, for the catchments in the individual calibration groups (Group 1–5) given by the climatic zones, and for all validation catchments for the calibration period (2000–2007) as well as the validation period (1983–1999) using the new version of the HBV model.

in the validation period of Ep and minimizes the calibration time consumption. The period 1983–1999 was used for validation.

Yang et al. (2018) applied a set of regionalized model parameters to multiple catchments with similar climate conditions in Norway and found satisfactory runoff simulations, which are comparable with the results using distance-based regionalization approaches (see Fig. 8 in their paper). Following their study, we selected the regionalization approach based on climatic similarity in this study, i.e. multiple catchments with similar climatic conditions were calibrated simultaneously and the calibrated parameters sets were transferred to other ungauged grid cells within the same climatic zone. It was further assumed that each climatic zone has similar precipitation correction factors.

We used four climatic indices to classify the study domain into five distinct climatic zones. The indices are mean annual precipitation, precipitation seasonality index (the ratio between the three consecutive wettest and driest months), mean annual temperature and temperature seasonality index (the mean temperature of the hottest month minus the mean temperature of the coldest month). The indices were calculated for each $1 \times 1\,km^2$ grid cell and the K-Mean clustering method (Hartigan and Wong, 1979) was applied to the standardized indices for all modelling grid cells. Fig. 2 illustrates the five climatic zones after classification and Fig. 3 shows that the zones represent not only distinct climatic regimes but also different geophysical and land cover characteristics, though only the climate indices were used for classification. Zone 1 is mainly located in the far north region, with lowest temperature and precipitation but a high seasonality index. Zone 2 spreads in the mountainous region from the south to the north, so it has the highest elevation among the five zones. This zone is mainly covered by bedrock and heathers and has almost no forest. Zone 3 is the warmest and wettest area and is situated along the west coast. It has a low climate seasonality index and it is covered by both forest and heather. Zone 4 is mainly located in east Norway and zone 5 in the southern and central part of Norway. Both zones have high coverage of forest but they differ in all climatic indices moderately.

For calibration, we selected catchments with drainage areas predominantly situated within one distinct climate zone (Fig. 2). As a result, there are five calibration groups with 3, 20, 31, 13 and 18 catchments selected for zone 1–5, respectively. The remaining 39 catchments, which cover more than one climatic zone, form the validation group used for testing the spatial transferability of the model. We calibrated the model for each climatic zone in parallel to speed up the calibration procedure and improve the simulation results on discharge for the selected catchments.

Three criteria were used for multi-criteria optimization, the non-dimensional performance criterion of Nash and Sutcliffe (1970) efficiency (NSE), the percent bias (PBIAS) in water balance and the percent volume bias in the high-flow segment of the flow duration curve (ΔFHV, 0–0.02 flow exceedance probabilities) (Yilmaz et al., 2008). ΔFHV is an index for high flows, aiming to improve the simulation of floods. The parameter estimation routine PEST (Doherty and Skahill, 2006) was applied for finding the parameter set giving the best model performance using the multi-criteria defined above. Since PEST minimizes the difference between the criteria results and their ideal values (1 for NSE and 0 for biases), the calibration objective function containing three criteria at multiple gauges can be formulated as Eq. (3).

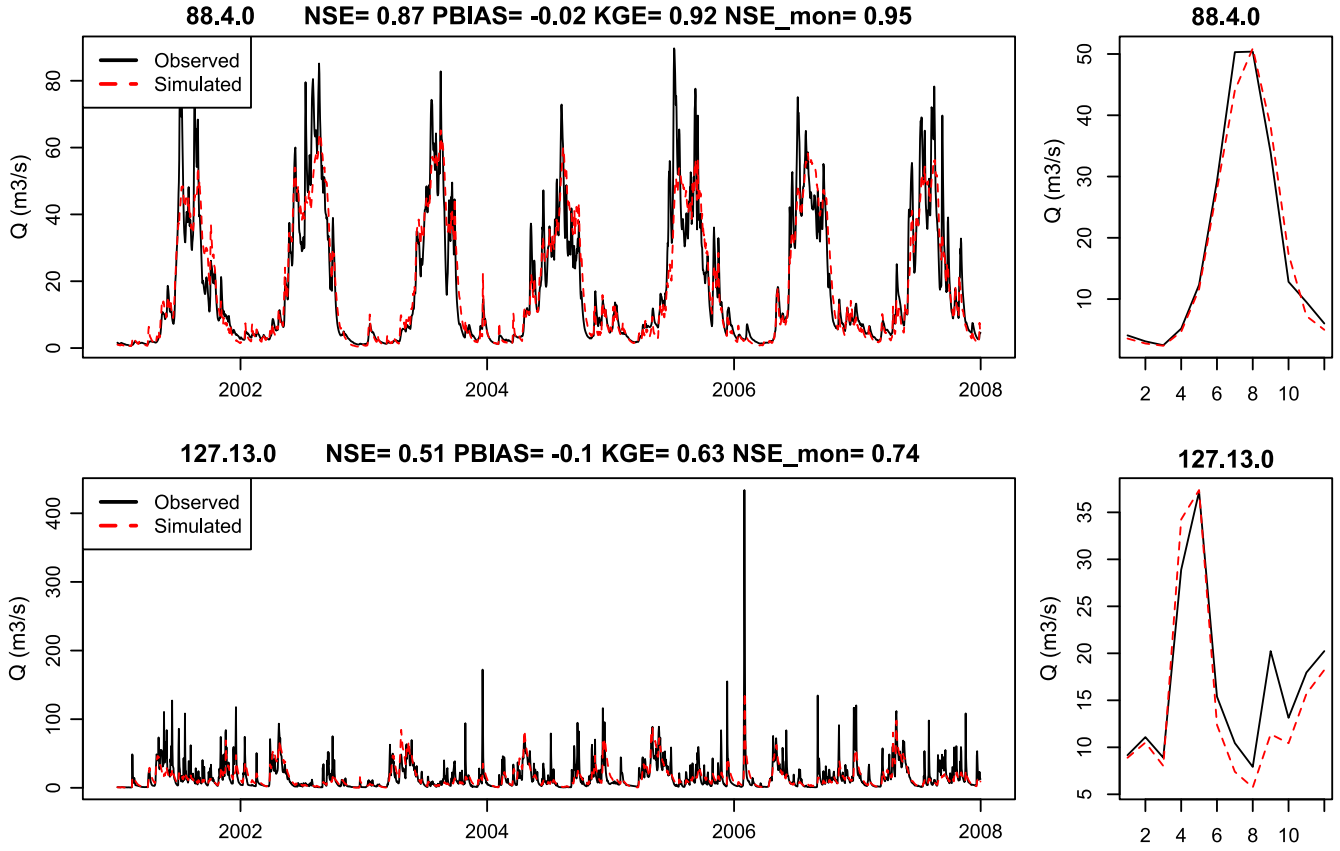


Fig. 7. Comparison of observed and simulated daily discharge (left sub-figures) at two selected gauges 88.4.0 and 127.13.0. The comparison of average monthly discharges is shown in the right panels.

$$\theta = W_{NSE} * \sum_{i=1}^{n} (1 - NSE_i)^2 + W_{PBIAS} * \sum_{i=1}^{n} (PBIAS_i)^2 + W_{\Delta FHV} * \sum_{i=1}^{n} (\Delta FHV_i)^2 \quad (3)$$

where W are weights and n are the number of calibration catchments within each group. There is an equal weight between catchments but the W_{NSE} is higher than the bias weights to achieve a good calibration performance. As the calibration time is significantly reduced by parallel calibration, 8–10 PEST runs were carried out with various initial parameter values, which were randomly generated within the parameter ranges listed in Table 1.

After the calibration, the model performance was evaluated using the daily NSE, PBIAS, Kling-Gupta Efficiency (KGE) (Kling et al., 2012) and monthly NSE for all 124 catchments in both calibration and validation periods.

3.3. Model input data

To setup the distributed version of the HBV model, we used a digital elevation model (DEM), soil and land cover raster maps with 1 km horizontal resolution. The DEM map is obtained from the Norwegian Mapping Authority, the Swedish National Land Survey and the National Land Survey of Finland. The soil map combines the information from the Geological Survey of Norway and the International Soil Reference and Information Centre (ISRIC) (http://isric.org). Eight soil types were reclassified for the modelling domain.

High-resolution land cover information was derived from the National Land Resource Map (Ahlstrøm et al., 2014) and a remote sensing based forest map (SAT-SKOG, Gjertsen and Nilsen, 2012), complemented by Corine Land Cover 2000 (https://www.eea.europa.eu/data-and-maps/data/clc-2000-raster-4) for areas outside Norway, and was aggregated to percentage share of different land cover types at $1 \times 1\,km^2$ HBV modelling resolution. The land cover classification distinguishes eight general land use types (open area, bog, built-up, cropland, heather, bedrock, lake, permanent ice and snow) and 12 structural forest types to better reflect spatial variability in hydrologically relevant land surface properties within forested areas. The forest types are based on the classification scheme developed by Majasalmi et al. (2018) for Fennoscandian forests which consists of three species groups (spruce, pine, and deciduous dominated) with four structural subgroups each. Look-up tables provided by Majasalmi et al. (2018) were used to parameterize maximum leaf area index and vegetation height in HBV, shortwave albedo was parametrized with values given in Bright et al. (2018).

The new version of the distributed HBV model requires seven climate variables with daily time steps as input: mean, maximum and minimum air temperature (Tmean, Tmax, Tmin), precipitation (P), shortwave incoming radiation (SWR), vapour pressure (VP) and wind speed. The seNorge2 precipitation dataset (Lussana et al., 2018) provides an observational gridded dataset over Norway from 1957 to the present day. This dataset is the most updated precipitation data from the Norwegian Meteorological Institute. It shows more precise estimates of annual average precipitation than seNorge version 1.1 (Mohr, 2008), which was used as input in the previous HBV application (Beldring et al., 2003). However, the product substantially underestimates precipitation in data-sparse areas and inherits the undercatch problem from the measurements (Lussana et al., 2018). The Tmax, Tmin and wind datasets are based on different post-processing procedures applied to NOrwegian ReAnalysis 10 km (NORA10) product (Reistad et al., 2011), which is a dataset of hindcasts (i.e. dynamical downscaling of a global reanalysis or analysis system over the Scandinavian peninsula). Tmin and Tmax combine the NORA10-derived daily extreme temperature fields with in-situ observations from a climate

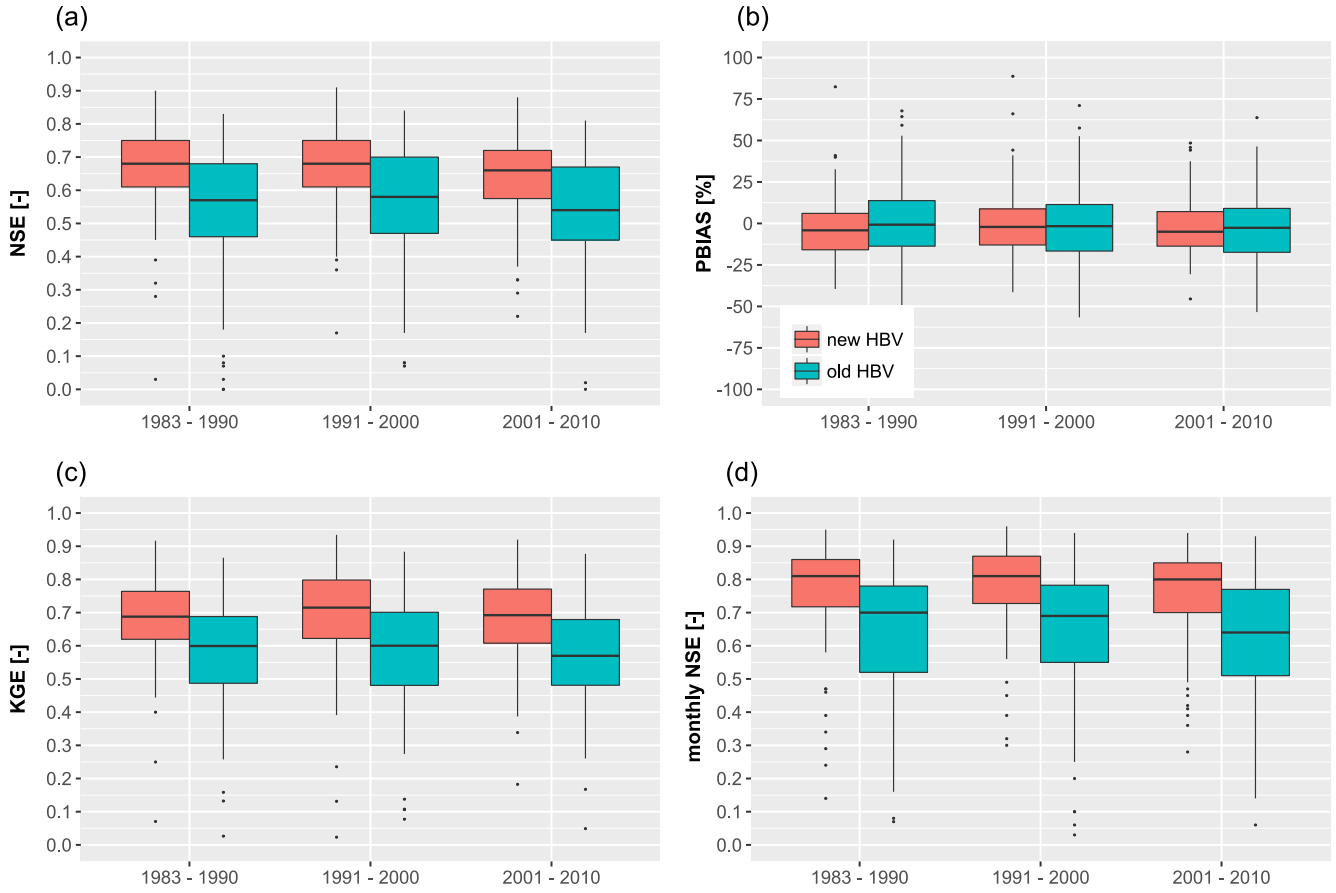


Fig. 8. Difference in model performance between the new and old version of the HBV model for all 124 catchments for three distinct periods.

station network from 1981 to 2012 (Lussana, 2017). The wind dataset is the result of a quantile mapping scheme aimed at adjusting the NORA10 bias to better match the climatology of a high-resolution numerical weather prediction model, such as Applications of Research to Operations at Mesoscale model for Meteorological Cooperation on Operational Numerical Weather Prediction (AROME – MetCoOp) (Mueller et al., 2017). All datasets have the same $1 \times 1\,km^2$ grid domain as the seNorge datasets, which covers the Norwegian mainland plus a strip of land extending into Sweden and Finland. Based on the Tmax, Tmin, P and wind data for the period 1982–2012, we applied the Mountain Microclimate Simulation Model (MTCLIM) (Bohn et al., 2013), which is embedded in the Variable Infiltration Capacity (VIC) hydrological model (Liang et al., 1994) (version 4.1.2), to generate remaining climate variables required by HBV (Tmean, SWR and VP) for the same period.

3.4. Data for potential evapotranspiration and snow water equivalent comparison

Ground measurements of Ep are scarce in Norway. There are only pan measurements from the "Thorsrud 2500" evaporimeter available at 42 locations across the country from 1967 to 1972 for the warm months (May–September) (Hetager and Lystad, 1974). Due to the cold climate and low winter insolation in Norway, the majority of E occurs in these warm months as shown in other high latitude areas (Shutov et al., 2006).

Two recently published satellite-based products were selected for comparison with the HBV estimates and the pan measurements. These are the MOD16 products of mean annual Ep from 2000 to 2012 (http://files.ntsg.umt.edu/data/NTSG_Products/MOD16/MOD16A2_MONTHLY.MERRA_GMAO_1kmALB/GEOTIFF_0.05degree/) and the GLEAM annual Ep available from 1980 to 2014 (https://www.gleam.eu/). The MOD16 products use a Penman-Monteith based method but estimate potential soil evaporation, soil evaporation from wet soil surface, canopy evaporation and potential transpiration from plant tissues separately (Mu et al., 2011). The climate forcing of MOD16 is derived from the daily global metrological reanalysis data from NASA's Global Modeling and Assimilation Office (GMAO) at $1 \times 1.25^{\circ}$ resolution. In contrast, the GLEAM Ep was calculated using the Priestley and Taylor equation (Miralles et al., 2011) and the climate forcing data combines the ERA-Interim and MSWEP v1.0 reanalysis data at 0.25° resolution.

None of the satellite products and the HBV simulation can be directly compared with the measurements due to different time spans. Therefore, we applied the VIC hydrological model to simulate the long-term Ep from 1961 to 2012 to bridge the temporal gaps. The VIC model gives an approximation of the recent warming impacts on Ep between the period 2000–2013 and 1967–1972 and helps to justify the comparison between the observations and estimations. The VIC model used the global VIC input parameters at 0.5° resolution (Nijssen et al., 2001) and the reanalysis climate forcing data from the Water and Global Change (WATCH) project (Weedon et al., 2011) as input. The Ep is estimated from an area-weighted sum of potential transpiration and potential soil evaporation using the Penman-Monteith based method for each grid cell.

The snow water equivalent (SWE) data (Saloranta, 2012) was calculated based on the measurements of snow depth and density recorded at stations operated by various hydropower companies since 1914. The stations were mainly located in the central mountain area in the southern, eastern and western Norway (see Appendix C). Most of the measurements (60%) are taken once per snow season around the time of the maximum annual SWE (March–April), but at many stations

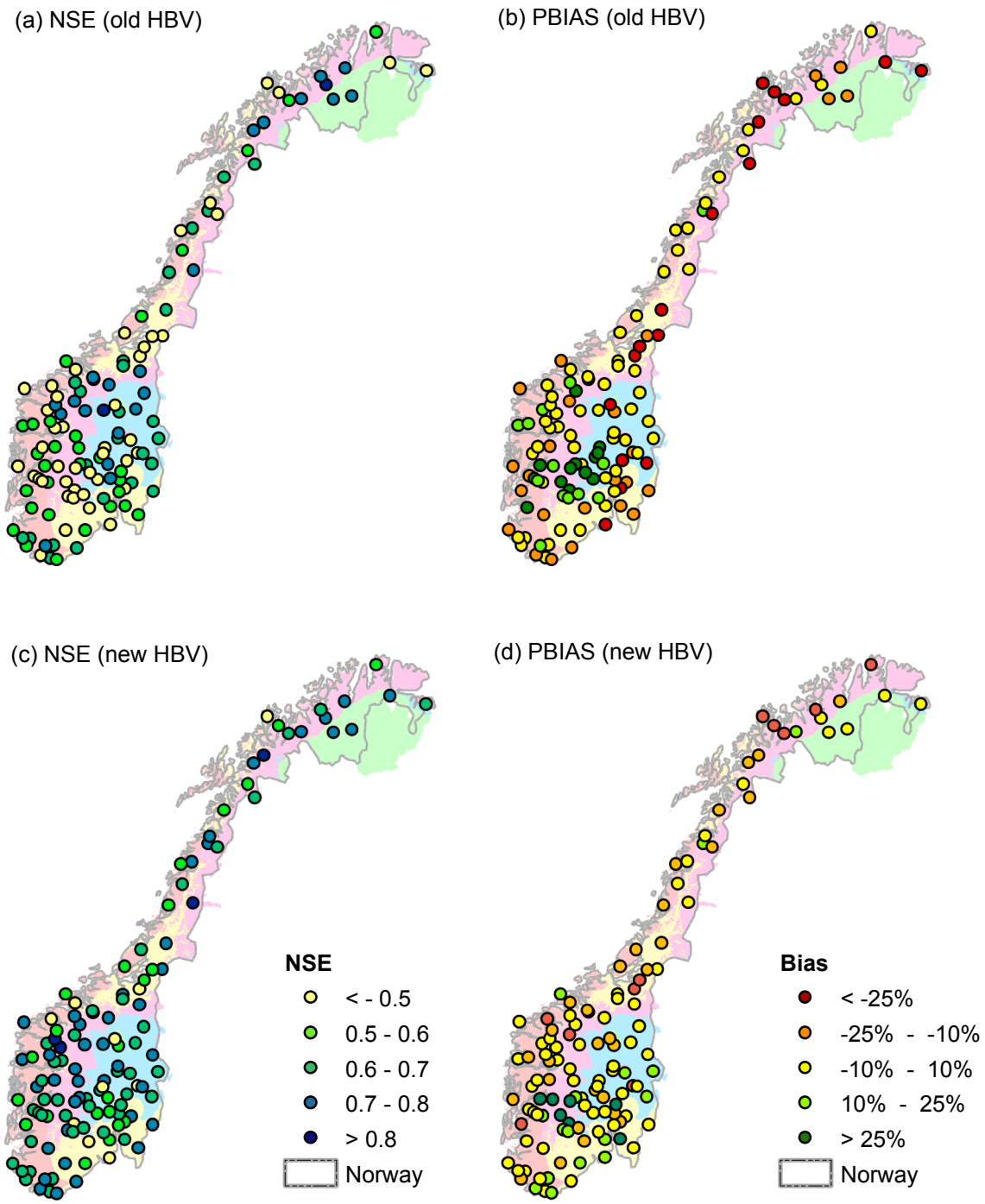


Fig. 9. Performance of the old (a and b) and new (c and d) version of the HBV model for the period 1983–2012. The background colours show the five climatic zones used in this study.

several (up to 13) measurements per snow season have been taken. The whole data set has been quality controlled by removing or correcting bad or duplicate values and outliers. Only data in the period 1983–2012 were used in model evaluation, resulting in 22,032 observations at 1151 stations (see Appendix C).

4. Results

4.1. Validation of potential evapotranspiration estimates

The Ep estimates for the summer months (May–Sep.) from the HBV model, MOD16, GLEAM and the VIC model are compared with the pan measurements in Fig. 4. The measurements are only available during the period from 1967 to 1972 in Norway while the model estimates span from 2000 to 2012. As we mentioned in Section 3.4, we ran the VIC model for the past 50 years to evaluate how Ep varied in different historical time slices. The long-term simulation with the VIC model shows that there is a slight increase of 6 mm/year in Ep from the period 1962–1970 to 2000–2012 (see Appendix D). This increase is very small compared to the average annual Ep that varies between 237 and 254 mm/year for the period spanning the past 50 years. Therefore, we assume that the difference of measured Ep between 2000 and 2012 and 1967–1972 is also negligible and we can directly compare the estimates with the measurements. As shown in Fig. 4, all estimates have good

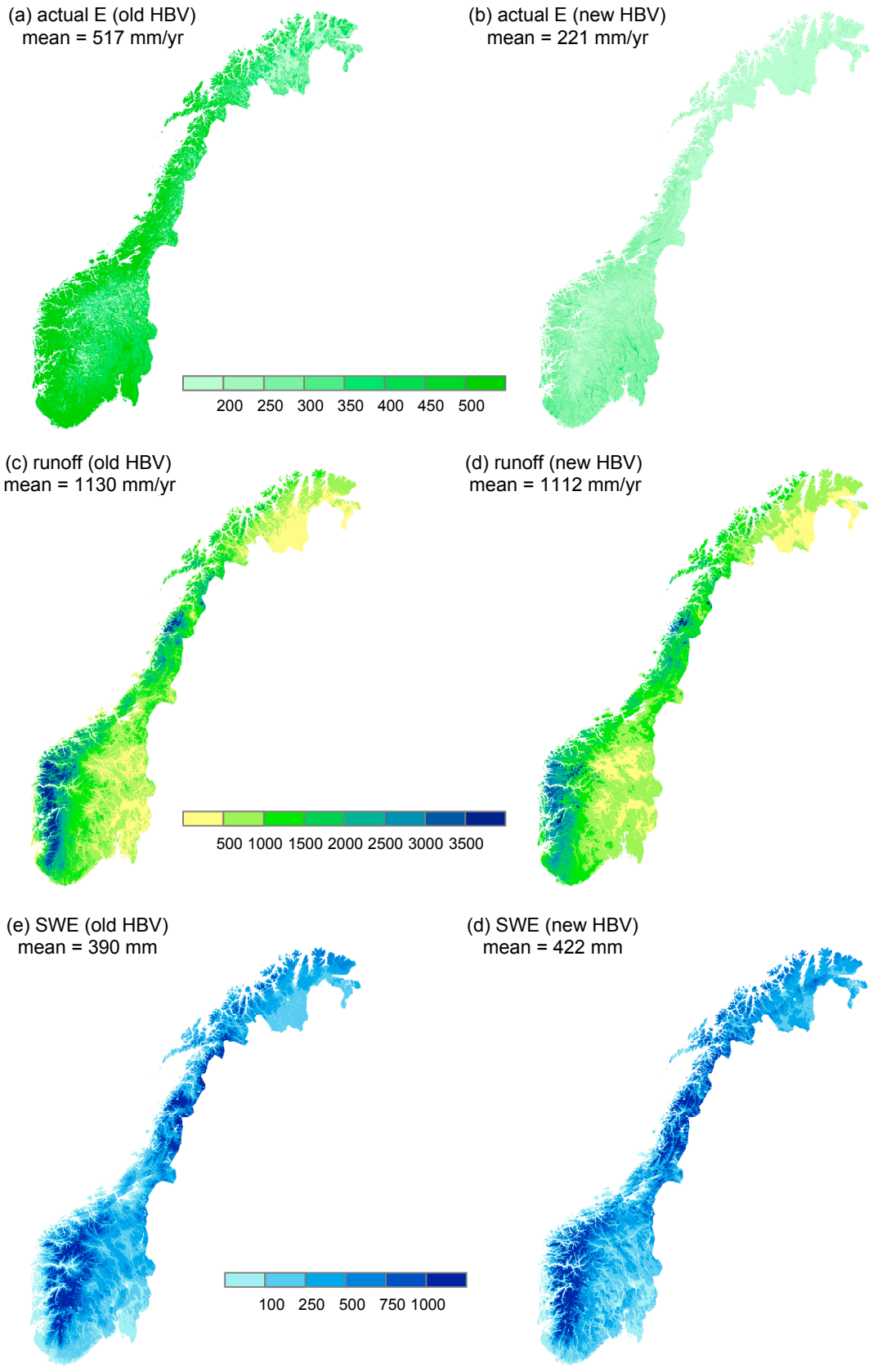


Fig. 10. Average annual actual evapotranspiration (E) (a and b) and runoff (c and d), as well as maximum snow water equivalent (SWE) (e and f) in 1983–2012 using the old and new version of HBV.

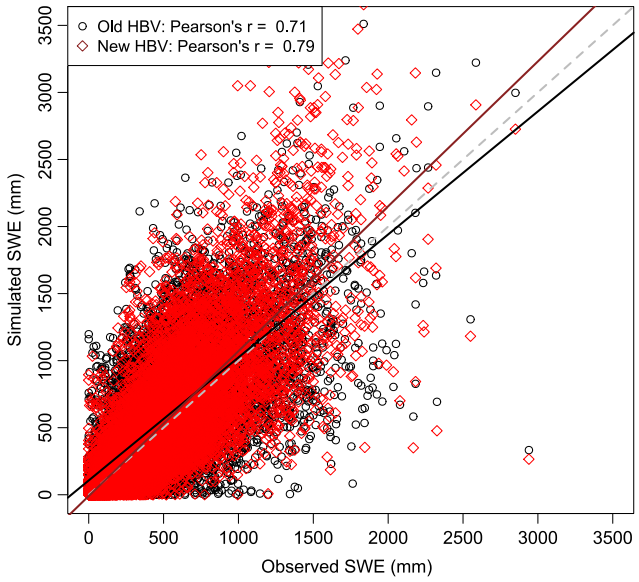


Fig. 11. Comparison between the observed snow water equivalent (SWE) at 1151 stations and the simulated SWE using the old and new version of HBV in 1983–2012. Grey dash line: 1:1 line; dark red line: linear regression line to the new HBV results; black line: linear regression line to the old HBV results.

agreement with an average annual Ep of 230–260 mm/year except the MOD16 product. The MOD16 shows a very high average annual Ep of more than 400 mm/year. The measurements indeed suggest that the Ep is significantly overestimated by the MOD16 especially for the Finnmark region, where the Ep is below 200 mm/year. In contrast, the HBV, GLEAM and VIC models give reasonable estimates compared to the measurements, with average annual Ep ranging from < 200 to 300 mm/year in the north and Finnmark and from 200 to > 500 mm/year in other regions. In addition, the HBV model presents the finest resolution among all products so that it can clearly reflect the spatial distribution of different land cover types and climate regimes.

Fig. 5 compares the estimated Ep with the measurements at the 42 stations. The MODIS data gives the overall highest estimate and it is much higher than the measurements. The GLEAM model using the Priestley and Taylor equation shows slightly higher estimates with less variability with respect to the measurements. Both the HBV and VIC models, which consider vegetation characteristics, give slightly lower estimates with respect to the measurements. However, we should notice that the measurements at a point scale cannot be identical to the estimates for cells, which constitute various land use types. In addition, the pan measurements are considered as the maximum value of E from large lake surfaces (Hetager and Lystad, 1974), so they should in principle be larger than the estimates using the Penman-Monteith approaches. Therefore, based on the assumption that the Ep varied only marginally in the past 50 years indicated by the VIC results, we conclude that the Penman-Monteith approach with prescribed vegetation parameters seems to provide reasonable Ep estimates for the mainland of Norway compared to the measurements.

4.2. Calibration and validation results on river discharge

The HBV model was calibrated against discharge observations from 85 catchments, which were classified into 5 groups based on the climatic zones (see Section 3.2 for details). For each group of catchments, the parameters listed in Table 1 were estimated to minimize the objective function shown in Eq. (3). As a result, there are 5 sets of parameters for the HBV model to simulate the whole land surface of Norway. Fig. 6 shows the evaluation results for different catchment groups within the calibration period 2000–2007. In general, the HBV model can reproduce the daily discharge well for most calibration catchments, with a median daily NSE of 0.68, PBIAS of -3%, KGE of 0.70 and monthly NSE of 0.80. The model shows different performance between the catchment groups and the median daily NSE/KGE varies from 0.78/0.83 for group 1 to 0.56/0.64 for group 5. However, we should note that there are only three catchments in group 1, so we cannot conclude that the HBV can reproduce the discharge in Finnmark better than in other regions. Different from the daily NSE and KGE criteria, the median PBIAS of all five groups are close to zero and about half of the catchments have a volume error within $\pm$ 12.5%. The monthly NSE shows lower variability between the groups than the daily values, and is above 0.7 for more than 75% of the calibration catchments in all groups except group 3.

Fig. 7 shows two examples of comparison between the observed and simulated daily/monthly discharge and explains why the daily and monthly NSEs differ significantly. The first example is the gauge 88.4, which is a glacier covered catchment in climatic zone 1 and shows the best simulation performance among all catchments. The HBV model reproduces the distinct seasonality of discharge well with both daily and monthly time steps. The second example is the gauge 127.13, which is located in Trøndelag and shows one of the poorest simulation results. The HBV model underestimated most discharge peaks significantly at the daily time step. However, the model reproduced the monthly peak satisfactorily. The results show that the model is not capable to reproduce flood peaks in some catchments even though the high flow index was included in the calibration objective function. It also indicates that the monthly simulation results are more reliable than the daily ones for such large scale modelling.

Fig. 6 further demonstrates a good temporal and spatial transferability of the HBV model. In general, the model presents a good performance for all calibrated catchments in the validation period, with median daily NSE of 0.67, PBIAS of -1%, KGE of 0.69 and monthly NSE of 0.78. The daily NSE and KGE show slightly better model performance in terms of the lower quantile and the smallest values in the boxplots, while the box of PBIAS expands by 2–3%, indicating larger biases in some catchments. The model performance declined for group 1 shown by the NSE, PBIAS and monthly NSE while the simulated discharge was improved for group 5 with the daily time step. For the spatial validation catchments, all criteria show a similar or better model performance compared to the calibration ones in both periods. This indicates that the calibrated parameterizations are appropriate for the non-gauged grid cells and the spatial distribution of simulated runoff should be reliable.

Finally, we compared the performance criteria using the new version of HBV with the ones using the old version for all 124 catchments (Beldring et al., 2003) in various periods in Fig. 8. The calibration period of the previous HBV application starts from 1991 to 2000. The comparison shows that the new methods improve the discharge simulation substantially for most catchments in all periods. The median values of the daily NSE, KGE and monthly NSE increase by 0.08–0.16. Though the median PBIAS of the previous simulation is also close to zero, the extent of the boxes and whiskers is wider than the one of the new simulation, indicating larger biases in the old than in the new simulations. However, both HBV versions show a robust temporal transferability for all catchments.

In addition to the boxplots, we compared the distribution of daily NSE and PBIAS in the whole simulation period (1983–2012) using the two HBV versions in Fig. 9. In general, there is no obvious spatial pattern of NSE in both simulation runs, but the spatial pattern of PBIAS varies in some regions. For example, the new HBV generated less bias in the east and Finnmark while the old one performs better in terms of bias along the coastline in Trøndelag and north regions. Significant positive biases (dark green dots) are found in Eastern and Western Norway for both simulations. The spatial pattern of PBIAS partly reflects the bias distribution of the seNorge precipitation data, which is a big challenge of climate input data at fine resolutions.

4.3. Hydrological components

The climate impacts on actual evapotranspiration, runoff and snow water equivalent (SWE) were assessed in the report "Climate in Norway 2100". In this paper, we compare these water components simulated by both versions of the HBV model to illustrate the differences in estimates due to model structures and calibration procedures (Fig. 10) for the whole simulation period (1983–2012). The results show distinct differences in actual evapotranspiration and a high level of agreement in runoff between the previous and current simulations. The old long-term average evapotranspiration estimate is more than double of the evapotranspiration estimated in this study and it is much higher than the measured Ep shown in Fig. 4. This is mainly caused by the overestimation of precipitation in the seNorge 1.1 precipitation data (Saloranta, 2014) and the calibration of Ep in the previous application (see Table 1).

Both HBV simulations show that the long-term average annual runoff is over 1100 mm for the mainland of Norway. This high agreement is not surprising because both calibrations are against discharge measurements from almost the same gauges. Nevertheless, the spatial distribution of the two estimates differs at regional scales. For example, the previous HBV simulation shows the annual runoff of more than 3500 mm in large parts of western Norway while the new simulation has moderate estimates in this region. It is hard to conclude that one spatial distribution of annual runoff outperforms the other because the distribution of PBIAS at gauges does not show a systematic improvement using the new version of HBV (Fig. 9). However, it should be in general more precise in the new simulation as shown in Fig. 8.

The long-term annual maximum SWE simulated by the two versions of HBV model differs by approximately 30 mm and the spatial distribution varies depending on regions. The new model simulated higher SWE especially in the north, where the model was calibrated separately for climate zone 1 and 2. Nevertheless, both models show distinct SWE depth between the western and eastern Norway. The simulated SWE from both HBV versions were also compared with observed SWE at 1151 stations in Norway (Fig. 11). The Pearson correlation coefficients indicate that the new HBV performs better in term of SWE than the old one in general. The regression lines show that the new HBV simulated SWE less than 500 mm very well, but overestimated high SWE values (> 1000 mm).

5. Discussion

There is an increasing trend of using satellite derived data to improve hydrological modelling especially for data scarce regions. However, different studies showed that the satellite-based data might not give reliable estimates in specific regions. For example, poor agreement between the MODIS16 actual evapotranspiration and the data from eddy covariance (EC) fluxes were found in South Africa (Ramoelo et al., 2014) and irrigation regions in China (Tang et al., 2015). Yang et al. (2017) also found that the GLEAM data significantly overestimated the EC measurements at four forest sites and one cropland site in China. The results from this study contribute to a better understanding of the suitability and the differences between Ep products in Norway. They show that the MODIS Ep may not be an appropriate product for Norway compared with other products. In contrast, the Penman-Monteith method using ground-based vegetation data can provide reasonable Ep estimates in the mainland of Norway at a fine resolution. Therefore, the satellite-based estimates should be carefully evaluated before using them in hydrological modelling either for calibration or as input data for regional studies. Finally, we should keep in mind that the measured and estimated Ep are from two different historical time slices in this study. The comparison results in Figs. 4 and 5 are strongly based on the assumption derived from the VIC long-term simulation, which is also highly determined by the quality of input data, such as the WATCH climate data and vegetation data. New measurements in future are strongly recommended for a more reliable validation of Ep estimates in Norway.

Using a more physically approach does not guarantee an improved model performance (Li et al., 2015; Magnusson et al., 2015; Orth et al., 2015; te Linde et al., 2008). However, the results in this study show a significant improvement on the spatial and temporal transferability of model parameterizations. This improvement can be attributed to either a more robust estimation of Ep or the use of a better regionalization approach, or a combination of both. The current approaches help to reduce the uncertainty of simulated hydrological components in Norway, especially for E. In addition, the distribution of performance criteria for monthly discharges in this study are comparable with the best results in Yang et al. (2018), who used a lumped conceptual model and a combined spatial proximity and physical similarity regionalization approach for multiple Norwegian catchments. Hence, this study presents a robust hydrological modelling for both river discharges and hydrological components at the national scale of Norway.

Finally, the undercatch problem in the seNorge 2 precipitation data (Lussana et al., 2018) remains a significant challenge for hydrological modelling in Norway. Although the precipitation undercatch can be corrected for each climatic zone, the large PBIAS at specific gauges requires further improvement of the precipitation data in terms of magnitude and spatial distribution. Ideally, the precipitation should not be corrected in the hydrological modelling and the hydrological model should still show satisfactory performance for all studied catchments. Hence, there is still a long-term work of improving the hydrological modelling with the up-to-date climate forcing data at fine resolutions. This can be achieved by either improving the interpolation method for observed precipitation data or applying a fully coupled land-atmosphere modelling system for data scarce regions as it is shown by Naabil et al. (2017).

6. Conclusion

This study presents the recent progress in improving the distributed version of HBV for the mainland of Norway. Firstly, we improved the input data to the model by collecting multi-source vegetation data at high spatial resolution and ground-based meteorological forcing data (e.g. maximum/minimum air temperature and wind speed). Secondly, we fed the high-resolution data to the Penman-Monteith equation to estimate potential evapotranspiration and validated the results by comparing with measurements and different satellite-based products. Thirdly, we applied regionalized parameterizations to calibrate the HBV model against discharge for multiple catchments. The calibration and validation results show a significant improvement on the simulated discharge as well as hydrological components, e.g. actual evapotranspiration, compared to previous simulation results. The transferability of model parameterizations was also enhanced yielding more robust simulation results both in time and in space. As a result, the current model setup is more appropriate for climate and land use impact studies and it is expected to project more reliable changes in discharge and hydrological components in Norway.

Further improvements are required for the existing national wide hydrological model. Improved precipitation data is urgently appreciated for a sound hydrological modelling. It is also interesting to test other model structures (e.g. snow and river routing schemes) with different complexities and regionalization methods. A robust snow modelling plays an essential role in the hydrological modelling for high latitudes. As was shown in this study, a physically based method may improve the hydrological simulation in Norway. Thus, we plan to test a physically based snow model for all grid cells in the mainland of Norway and substitute it for the parametric snow module currently implemented in HBV.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank Norwegian Research Council for funding the I:CAN (Impacts: Climate, anthroposphere and nature) project no. 243803 and our colleague Tuomo Mikael Saloranta for providing the observed snow data.

Appendix

See Appendices A and B
See Appendices C and D

Appendix A
List of the 124 catchments for calibration and validation.

Nr.	Gauge ID	Latitude (°)	Longitude (°)	Area (km^2)	Group
1	2.11	62.36	11.48	119.13	Validation
2	2.28	61.22	10.27	866.25	Group4
3	2.32	61.85	10.22	463.20	Validation
4	2.142	60.64	12.05	1646.00	Group4
5	2.145	61.33	10.28	11212.84	Validation
6	2.265	61.95	11.08	618.36	Validation
7	2.268	61.80	8.45	790.90	Validation
8	2.279	60.13	11.08	432.35	Group5
9	2.291	62.01	7.87	262.27	Group2
10	2.323	60.88	11.32	42.47	Group4
11	2.439	61.18	10.89	377.09	Validation
12	2.479	62.01	10.00	156.81	Validation
13	2.604	60.87	11.56	15451.76	Group4
14	2.614	61.86	9.40	1833.91	Validation
15	2.633	59.36	11.52	87.13	Group5
16	2.634	60.67	10.81	182.60	Group4
17	3.22	59.55	10.86	299.29	Group5
18	6.10	59.99	10.80	7.03	Group5
19	6.71	60.17	10.50	7.63	Group5
20	8.2	59.89	10.51	190.19	Group5
21	12.70	60.95	9.63	570.17	Group4
22	12.114	60.21	10.02	492.53	Group4
23	12.150	60.49	8.82	24.79	Group4
24	12.171	60.71	9.46	79.75	Group4
25	12.178	60.15	9.43	309.77	Group4
26	12.192	59.52	10.12	74.68	Group5
27	12.193	59.70	9.79	51.60	Group5
28	12.209	60.56	9.88	542.89	Group4
29	12.215	60.89	8.33	119.43	Group2
30	12.286	60.40	10.54	113.26	Group5
31	12.290	60.81	9.57	2976.54	Validation
32	15.49	60.49	8.48	59.13	Validation
33	15.53	60.31	9.01	94.07	Group4
34	15.79	60.38	8.26	1177.29	Group2
35	15.174	59.08	10.18	25.58	Group5
36	16.66	59.84	8.32	6.60	Group2
37	16.75	59.67	8.07	118.28	Validation
38	16.122	59.64	8.31	42.69	Validation
39	16.132	59.76	8.79	33.10	Group2
40	16.140	60.02	7.87	822.26	Group2
41	16.193	59.43	9.13	156.65	Group5
42	18.10	58.88	9.03	236.23	Group5
43	19.107	58.44	8.70	39.32	Group5
44	20.2	58.84	8.10	276.42	Group5
45	22.4	58.12	7.53	1757.70	Group3
46	22.16	58.45	7.58	182.20	Group3
47	22.22	58.09	7.84	203.58	Group3
48	24.8	58.16	7.07	121.41	Group3
49	24.9	58.40	7.22	272.16	Group3
50	25.24	59.16	7.21	96.82	Validation
51	26.20	58.54	6.50	77.25	Group3
52	26.29	58.35	6.34	52.95	Group3
53	27.24	58.53	6.15	184.72	Group3
54	28.7	58.69	5.65	139.35	Group3
55	35.16	59.28	6.38	45.34	Group3
56	36.9	59.95	6.98	45.87	Group2
57	41.1	59.68	6.01	130.61	Group3
58	46.9	60.07	6.28	7.01	Group3

(*continued on next page*)

Appendix A (*continued*)

Nr.	Gauge ID	Latitude (°)	Longitude (°)	Area (km^2)	Group
59	48.1	60.01	6.55	470.22	Group3
60	48.5	59.97	6.73	120.50	Group3
61	50.1	60.36	6.75	232.73	Validation
62	50.13	60.39	7.40	262.61	Group2
63	55.4	60.25	5.44	50.09	Group3
64	62.5	60.63	6.29	1092.04	Group3
65	62.10	60.80	6.50	158.28	Group3
66	72.5	60.85	7.11	268.16	Validation
67	73.27	61.17	8.10	30.32	Group2
68	75.23	61.35	7.40	45.92	Validation
69	76.5	61.67	7.24	65.29	Group2
70	76.5	61.67	7.24	65.29	Group2
71	77.3	61.30	7.01	110.93	Group3
72	78.8	61.45	6.74	40.42	Group3
73	82.4	61.25	5.39	218.96	Group3
74	83.2	61.33	5.89	508.13	Group3
75	87.10	61.72	6.52	217.12	Group3
76	88.4	61.86	6.89	234.88	Validation
77	91.2	62.16	5.17	25.72	Group3
78	97.1	62.33	6.59	88.95	Group3
79	98.4	62.07	6.93	138.78	Validation
80	103.1	62.28	8.12	435.51	Group2
81	103.40	62.47	7.78	1099.42	Validation
82	105.1	62.79	7.73	137.59	Group3
83	107.3	62.97	7.16	24.23	Group3
84	109.9	62.51	9.59	745.38	Validation
85	109.29	62.61	8.73	85.44	Validation
86	109.42	62.65	8.69	2437.41	Validation
87	122.9	63.11	10.23	3086.41	Validation
88	122.11	62.89	11.18	654.24	Validation
89	122.17	62.99	10.25	546.17	Validation
90	123.31	63.27	11.13	145.00	Group5
91	124.2	63.49	11.36	494.72	Group5
92	127.11	63.79	12.33	176.13	Validation
93	127.13	63.76	11.77	479.64	Group5
94	133.7	63.80	10.23	206.61	Group3
95	138.1	64.24	11.08	239.07	Group3
96	139.35	64.43	12.48	852.39	Validation
97	148.2	65.37	12.54	108.47	Group3
98	151.15	65.44	13.99	653.36	Group2
99	152.4	65.91	13.31	525.69	Validation
100	156.10	66.46	13.88	210.63	Group3
101	157.3	66.39	13.18	16.33	Group3
102	161.7	66.90	14.85	225.05	Group2
103	162.3	67.08	14.98	145.39	Validation
104	163.5	66.81	15.41	421.98	Group2
105	168.2	67.72	15.87	31.23	Validation
106	173.8	68.02	17.90	62.83	Group2
107	174.3	68.35	17.49	28.44	Group2
108	191.2	68.86	17.94	526.02	Validation
109	196.35	69.03	18.66	3110.65	Validation
110	200.4	69.98	19.10	136.01	Validation
111	203.2	69.73	19.86	92.48	Validation
112	206.3	69.52	20.53	200.48	Group2
113	208.3	69.53	21.38	1932.29	Validation
114	212.10	69.42	23.64	5620.81	Group1
115	212.27	69.80	23.24	624.18	Validation
116	212.49	70.03	22.94	144.89	Group2
117	223.2	70.14	24.76	877.14	Group2
118	230.1	70.88	27.51	19.31	Group2
119	234.13	69.43	24.94	2079.25	Group1
120	234.18	70.07	28.02	14161.40	Group1
121	247.3	69.66	30.38	128.89	Validation
122	311.6	61.26	12.32	4424.85	Group4
123	311.460	61.66	12.02	394.83	Validation
124	313.10	59.95	12.19	359.95	Group5

Appendix B

List of the prescribed parameter values used for the Penman-Monteith approach.

Land use type	Albedo[1] [−]	Vegetation height [m]	Leaf area index[2] [−]	Surface resistance[3] [s/m]	Tmin_open[3] [°C]	Tmin_close[3] [°C]	VPD_close[3] [Pa]	VPD_open[3] [Pa]	gl_sh[3] [m/s]	Cl[3] [m/s]	z0g[4] [m]
Open	0.25	0.2	2	40	12.02	−8	4200	650	0.02	0.007	0.001
Bog	0.17	0.5	2	50	8.8	−8	4400	650	0.04	0.0065	0.01
Build-up	0.11	10	4	0	12.02	−8	4200	650	0.02	0.007	0.02
Cropland	0.25	2	5	40	12.02	−8	4500	650	0.02	0.007	0.01
Heather	0.25	0.3	1	15	8.8	−8	4400	650	0.04	0.0065	0.02
Bedrock	0.12	0.1	0	0	12.02	−8	4200	650	0.02	0.007	0.001
Spruce 1	0.13	7.5	1.4	70	8.31	−8	3000	650	0.04	0.0032	0.02
Spruce 2	0.12	12.3	4.3	70	8.31	−8	3000	650	0.04	0.0032	0.02
Spruce 3	0.12	16.8	6.7	70	8.31	−8	3000	650	0.04	0.0032	0.02
Spruce 4	0.11	22	9.1	70	8.31	−8	3000	650	0.04	0.0032	0.02
Pine 1	0.15	7.5	0.9	70	8.31	−8	3000	650	0.04	0.0032	0.02
Pine 2	0.08	11.6	2.4	70	8.31	−8	3000	650	0.04	0.0032	0.02
Pine 3	0.1	17	2.3	70	8.31	−8	3000	650	0.04	0.0032	0.02
Pine 4	0.09	17.2	4.4	70	8.31	−8	3000	650	0.04	0.0032	0.02
Deciduou 1	0.16	4.9	0.5	100	9.09	−6	2900	650	0.04	0.0032	0.02
Deciduou 2	0.16	8.4	1.8	100	9.09	−6	2900	650	0.04	0.0032	0.02
Deciduou 3	0.14	12.2	3.9	100	9.09	−6	2900	650	0.04	0.0032	0.02
Deciduou 4	0.15	18.3	7	100	9.09	−6	2900	650	0.04	0.0032	0.02

Note: 1. Albedo is used to calculate the net incoming raditation to the land surface.

2. Leaf area index (LAI) varies for deciduous forest and crops seasonally and the values here are the maximum LAI.

3. Surface resistance (r_s) for vegetation are calculated using the parameters Tmin_open, Tmin_close, VPD_close, VPD_open, gl_sh and C_L. The values in this table are the maximum r_s.

Tmin_open is the air temperature at which there is no water stress on transpiration.

Tmin_close is the air temperature at which stomata close almost completely.

VPD_open is the vapor pressure deficit (VPD) value at which where is no water stress on transpiration.

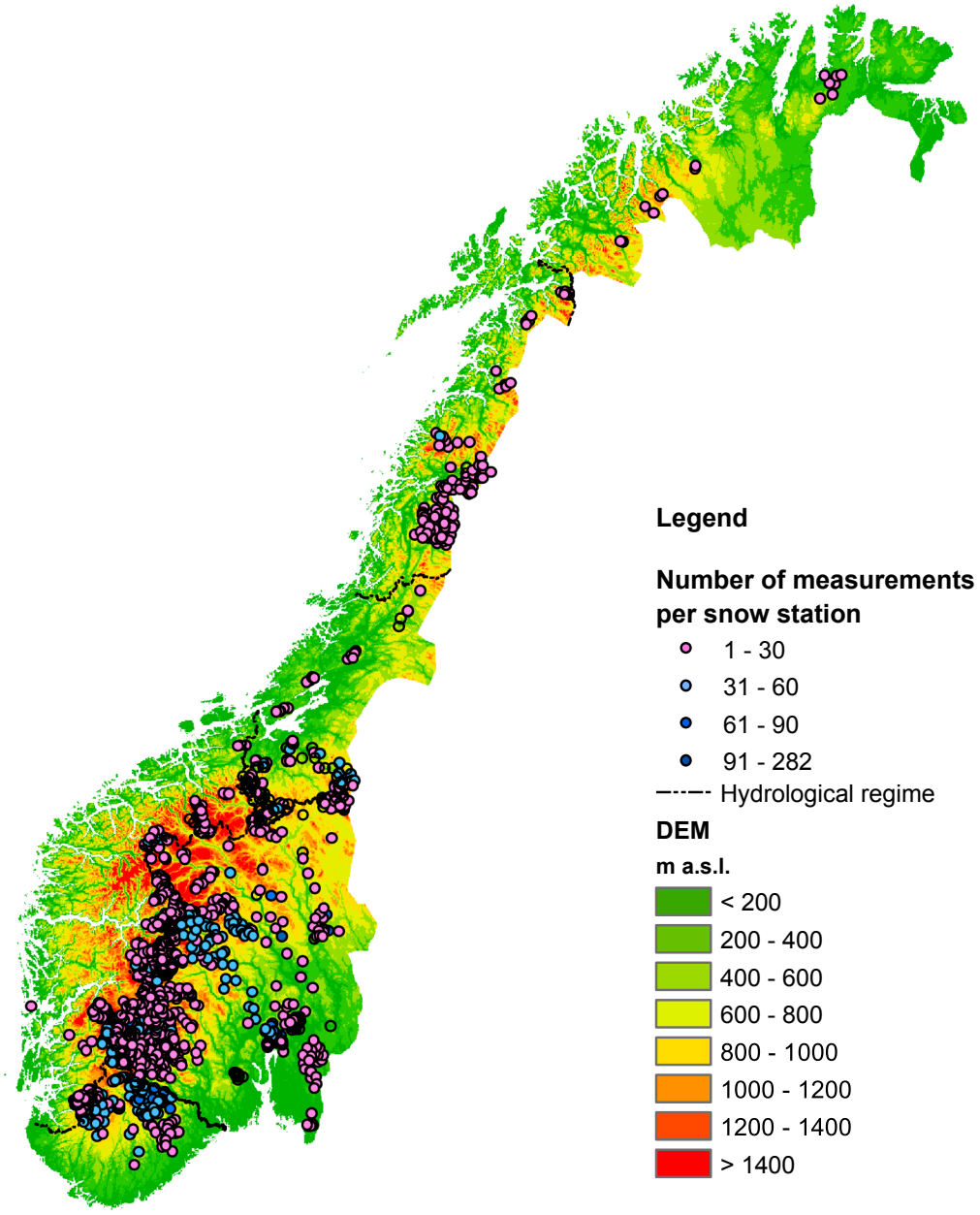
VPD_close is the VPD value at which stomata close almost completely.

gl_sh is leaf conductance to sensible heat per unit LAI.

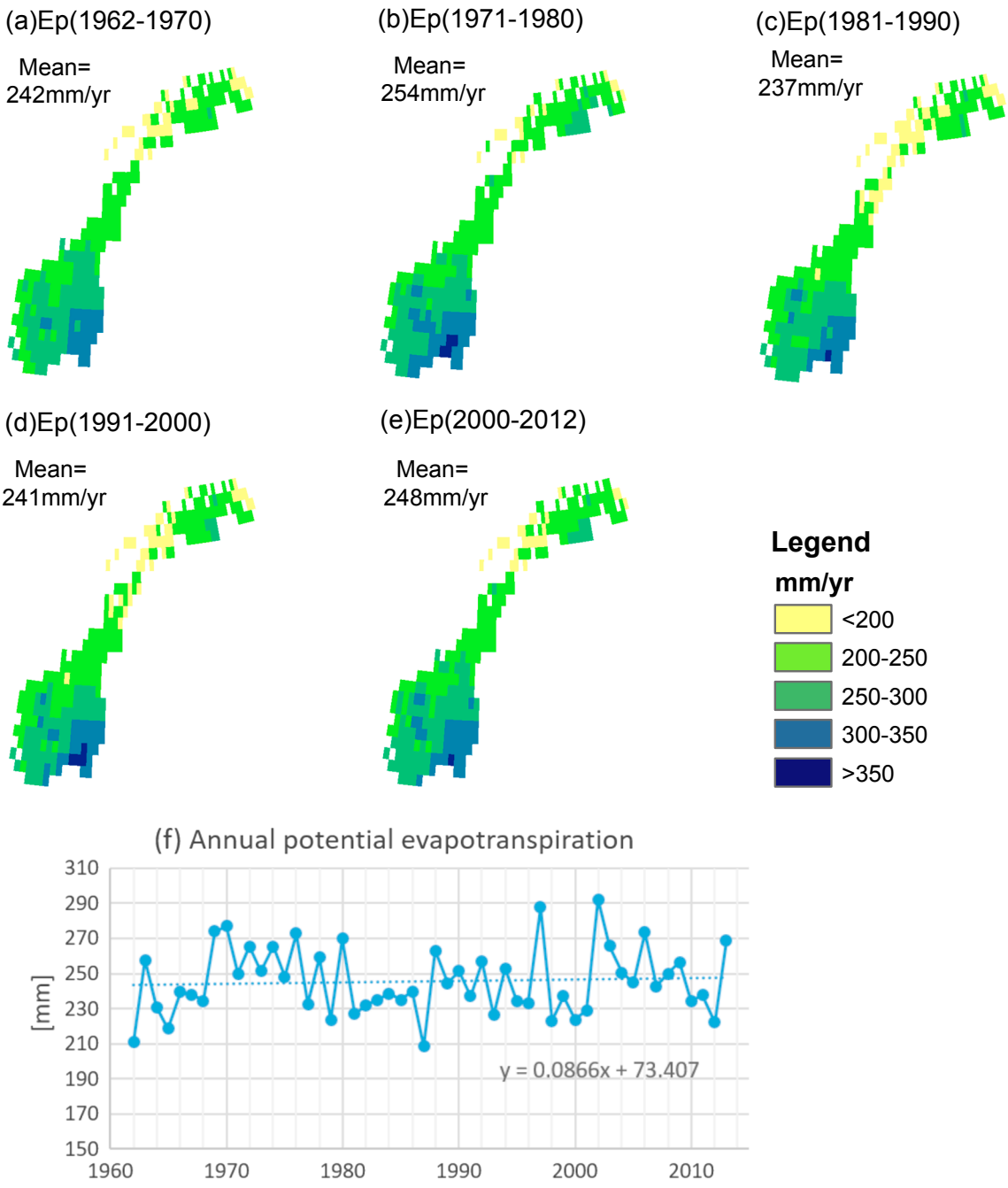
C_L is the mean potential stomatal conductance per unit leaf area

The details of equations and the paramters can be found in Mu et al. (2011) for vegetations.

4. z0g is the roughness parameter for ground, used in the estimation of aerodynamic resistances. More details see Federer et al. (1996).



Appendix C. Distribution of the snow stations operated by the hydropower companies.



Appendix D. Potential evapotranspiration (Ep) simulated by the Variable Infiltration Capacity model in different periods (a–e) and the annual Ep in each year (f).

References

Abbaspour, K.C., et al., 2015. A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. J Hydrol 524, 733–752. https://doi.org/10.1016/j.jhydrol.2015.03.027.

Ahlstrøm, A., Bjørkelo, K., Frydenlund, J., 2014. AR5 klassifikasjonssystem – Klassifikasjon av arealressurser. Norsk institutt for Skog og landskap, report nr. 6/2014, 38 pp. http://hdl.handle.net/11250/2440173.

Beck, H.E., et al., 2016. Global-scale regionalization of hydrologic model parameters. Water Resour. Res. 52 (5), 3599–3622. https://doi.org/10.1002/2015wr018247.

Beldring, S., Engeland, K., Roald, L.A., Saelthun, N.R., Vokso, A., 2003. Estimation of parameters in a distributed precipitation-runoff model for Norway. Hydrol. Earth Syst. Sci. 7 (3), 304–316. https://doi.org/10.5194/hess-7-304-2003.

Bergström, S., 1976. Development and Application of a Conceptual Runoff Model for Scandinavian Catchments. Swedish meteorological and hydrological institute, Norrköping, Sweden report nr. RHO 7.

Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources publications, pp. 443–476.

Bergström, S., 2006. Experience from applications of the HBV hydrological model from the perspective of prediction in ungauged basins. In: Andréassian, V., Hall, A., Chahinian, N., Schaake, J. (Eds.), Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiments. MOPEX. IAHS Publication, pp. 97–109.

Bohn, T.J., et al., 2013. Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. Agric. For. Meteorol. 176, 38–49. https://doi.org/10.1016/j.agrformet.2013.03.003.

Bowman, A.L., Franz, K.J., Hogue, T.S., 2017. Case studies of a MODIS-based potential evapotranspiration input to the sacramento soil moisture accounting model. J. Hydrometeorol. 18 (1), 151–158. https://doi.org/10.1175/Jhm-D-16-0214.1.

Bright, R.M., et al., 2018. Inferring surface albedo prediction error linked to forest structure at high latitudes. J. Geophys. Res. Atmos. 123 (10), 4910–4925. https://doi.org/10.1029/2018JD028293.

Coron, L., Andreassian, V., Perrin, C., Bourqui, M., Hendrickx, F., 2014. On the lack of robustness of hydrologic models regarding water balance simulation: a diagnostic approach applied to three models of increasing complexity on 20 mountainous catchments. Hydrol. Earth Syst. Sci. 18 (2), 727–746. https://doi.org/10.5194/hess-18-727-2014.

Demirel, M.C., et al., 2018. Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model. Hydrol. Earth Syst. Sci. 22 (2), 1299–1315. https://doi.org/10.5194/hess-22-1299-2018.
Doherty, J., Skahill, B.E., 2006. An advanced regularization methodology for use in watershed model calibration. J. Hydrol. 327 (3–4), 564–577. https://doi.org/10.1016/j.jhyrol.2005.11.058.
Donohue, R.J., McVicar, T.R., Roderick, M.L., 2010. Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. J. Hydrol. 386 (1–4), 186–197. https://doi.org/10.1016/j.jhydrol.2010.03.020.
Ershadi, A., McCabe, M.F., Evans, J.P., Wood, E.F., 2015. Impact of model structure and parameterization on Penman-Monteith type evaporation models. J. Hydrol. 525, 521–535. https://doi.org/10.1016/j.jhydrol.2015.04.008.
Federer, C.A., Voeroesmarty, C., Fekete, B., 1996. Intercomparison of methods for calculating potential evaporation in regional and global water balance models. Water Resour. Res. 32 (7), 2315–2321. https://doi.org/10.1029/96WR00801.
Fowler, K.J.A., Peel, M.C., Western, A.W., Zhang, L., Peterson, T.J., 2016. Simulating runoff under changing climatic conditions: revisiting an apparent deficiency of conceptual rainfall-runoff models. Water Resour. Res. 52 (3), 1820–1846. https://doi.org/10.1002/2015wr018068.
Førland, E.J. et al., 1996. Manual for operational correction of Nordic precipitation data, Norwegian Meteorological Institute, DNMI Klima Report, report nr. 24/1996, 72 pp.
Gjertsen, A.K., Nilsen, J.E., 2012. SAT-SKOG: Et skogkart basert på tolking av satellittbilder. Norsk institutt for Skog og landskap, report nr. 23/2012, 54 pp. http://hdl.handle.net/11250/2453917.
Hanssen-Bauer, I. et al., 2017. Climate in Norway 2100 - a knowledge base for climate adaptation. Norwegian Environment Agency, report nr. 1/2017, 48 pp.
Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a K-means clustering algorithm. Appl. Stat. 28, 100–108. https://doi.org/10.2307/2346830.
Hetager, S.E., Lystad, S.L., 1974. Evaporation from a free water surface – values based on measurements in the period 1967 – 1972. The Norwegian committee for the international hydrological decade programme, Oslo.
Hundecha, Y., Arheimer, B., Donnelly, C., Pechlivanidis, I., 2016. A regional parameter estimation scheme for a pan-European multi-basin model. J. Hydrol. –Reg. Stud. 6, 90–111. https://doi.org/10.1016/j.ejrh.2016.04.002.
Kaspersen, P.S., Halsnæs, K., Gregg, J.S., Drews, M., 2012. Methodological framework, analytical tool and database for the assessment of climate change impacts, adaptation and vulnerability in Denmark. DTU Management Engineering Report; No. 11.2012.
Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol. 424, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011.
Li, H., Xu, C.Y., Beldring, S., 2015. How much can we gain with increasing model complexity with the same model concepts? J. Hydrol. 527, 858–871. https://doi.org/10.1016/j.jhydrol.2015.05.044.
Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99 (D7), 14415–14428. https://doi.org/10.1029/94JD00483.
Lopez, P.L., Sutanudjaja, E.H., Schellekens, J., Sterk, G., Bierkens, M.F.P., 2017. Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. Hydrol. Earth Syst. Sci. 21 (6), 3125–3144. https://doi.org/10.5194/hess-21-3125-2017.
Lussana, C., 2017. Spatial Interpolation of Daily Minimum, Maximum and Mean Temperature. Norwegian Meteorological Institute, pp. 52 report nr. 02/2017.
Lussana, C., et al., 2018. seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day. Earth Syst. Sci. Data 10 (1), 235–249. https://doi.org/10.5194/essd-10-235-2018.
Magnusson, J., et al., 2015. Evaluating snow models with varying process representations for hydrological applications. Water Resour. Res. 51 (4), 2707–2723. 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.
McAfee, S.A., 2013. Methodological differences in projected potential evapotranspiration. Clim. Change 120 (4), 915–930. https://doi.org/10.1007/s10584-013-0864-7.
Merz, R., Parajka, J., Bloschl, G., 2011. Time stability of catchment model parameters: Implications for climate impact analyses. Water Resour. Res. 47. https://doi.org/10.1029/2010wr009505.
Miralles, D.G., et al., 2011. Global land-surface evaporation estimated from satellite-based observations. Hydrol. Earth Syst. Sci. 15 (2), 453–469. https://doi.org/10.5194/hess-15-453-2011.
Miralles, D.G., et al., 2016. The WACMOS-ET project – Part 2: evaluation of global terrestrial evaporation data sets. Hydrol. Earth Syst. Sci. 20 (2), 823–842. https://doi.org/10.5194/hess-20-823-2016.
Mohr, M., 2008. New routines for gridding of temperature and precipitation observations for “seNorge. no”, the Norwegian meteorological Institute, report nr. 08/2008, 43 pp.
Monteith, J.L., 1965. Evaporation and environment. Symp. Soc. Exp. Biol. 19, 205–234.
Mu, Q.Z., Zhao, M.S., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. 115 (8), 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019.
Mueller, M., et al., 2017. AROME-MetCoOp: a nordic convective-scale operational weather prediction model. Weather For. 32 (2), 609–627. https://doi.org/10.1175/Waf-D-16-0099.1.
Naabil, E., Lamptey, B.L., Arnault, J., Kunstmann, H., Olufayo, A., 2017. Water resources management using WRF-Hydro modelling system: case study of the Tono basin in West Africa. J. Hydro Regional Stud. 12, 196–209. https://doi.org/10.1016/j.ejrh.2017.05.010.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I -A discussion of principles. J. Hydrol. 10 (3), 282–290.
Nijssen, B., O'Donnell, G.M., Lettenmaier, D.P., Lohmann, D., Wood, E.F., 2001. Predicting the discharge of global rivers. J Clim. 14 (15), 3307–3323. https://doi.org/10.1175/1520-0442(2001) 014<3307:Ptdogr>2.0.Co;2.
Olsson, C., Jonsson, A.M., 2014. Process-based models not always better than empirical models for simulating budburst of Norway spruce and birch in Europe. Global Change Biol. 20 (11), 3492–3507. https://doi.org/10.1111/gcb.12593.
Olsson, J., et al., 2016. Hydrological climate change impact assessment at small and large scales: key messages from recent progress in Sweden. Climate 4 (3). https://doi.org/10.3390/cli4030039.
Orth, R., Staudinger, M., Seneviratne, S.I., Seibert, J., Zappa, M., 2015. Does model performance improve with complexity? A case study with three hydrological models. J. Hydrol. 523, 147–159. https://doi.org/10.1016/j.jhydrol.2015.01.044.
Oudin, L., Andreassian, V., Perrin, C., Michel, C., Le Moine, N., 2008. Spatial proximity, physical similarity, regression and ungaged catchments: a comparison of regionalization approaches based on 913 French catchments. Water Resour Res 44 (3). https://doi.org/10.1029/2007WR006240.
Parajka, J., Bloschl, G., Merz, R., 2007. Regional calibration of catchment models: Potential for ungauged catchments. Water Resour. Res. 43 (6). https://doi.org/10.1029/2006wr005271.
Rakovec, O., et al., 2016. Multiscale and multivariate evaluation of water fluxes and states over European river basins. J Hydrometeorol 17 (1), 287–307. https://doi.org/10.1175/Jhm-D-15-0054.1.
Ramoelo, A., et al., 2014. Validation of global evapotranspiration product (MOD16) using flux tower data in the african savanna South Africa. Remote Sens.-Basel 6 (8), 7406–7423. https://doi.org/10.3390/rs6087406.
Reistad, M., et al., 2011. A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. J. Geophys. Res.-Oceans 116. https://doi.org/10.1029/2010jc006402.
Saloranta, T., 2014. New version (v.1.1.1) of the seNorge snow model and snow maps for Norway. 6 – 2014, Norwegian Water Resources and Energy Directorate, Oslo, Norway, report nr. 6/2014, 36 pp.
Saloranta, T.M., 2012. Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model. Cryosphere 6 (6), 1323–1337. https://doi.org/10.5194/tc-6-1323-2012.
Seiller, G., Anctil, F., 2016. How do potential evapotranspiration formulas influence hydrological projections? Hydrol. Sci. J. 61 (12), 2249–2266. https://doi.org/10.1080/02626667.2015.1100302.
Shutov, V., Gieck, R.E., Hinzman, L.D., Kane, D.L., 2006. Evaporation from land surface in high latitude areas: a review of methods and study results. Nord Hydrol. 37 (4–5), 393–411. https://doi.org/10.2166/nh.2006.022.
Shuttleworth, W.J., Wallace, J.S., 1985. Evaporation from sparse crops-an energy combination theory. Quart. J. Roy. Meteorol. Soc. 111 (469), 839–855. https://doi.org/10.1002/qj.49711146910.
Spies, R.R., Franz, K.J., Hogue, T.S., Bowman, A.L., 2015. Distributed hydrologic modeling using satellite-derived potential evapotranspiration. J. Hydrometeorol. 16 (1), 129–146. https://doi.org/10.1175/Jhm-D-14-0047.1.
Tallaksen, L.M., Schunselaar, S., van Veen, R., 1996. Comparative model estimates of interception loss in a coniferous forest stand. Nordic Hydrol. 27, 143–160. https://doi.org/10.2166/nh.1996.0001.
Tang, R.L., et al., 2015. Multiscale validation of the 8-day MOD16 evapotranspiration product using flux data collected in china. IEEE J-Stars 8 (4), 1478–1486. https://doi.org/10.1109/Jstars.2015.2420105.
te Linde, A.H., Aerts, J.C.J.H., Hurkmans, R.T.W.L., Eberle, M., 2008. Comparing model performance of two rainfall-runoff models in the Rhine basin using different atmospheric forcing data sets. Hydrol. Earth Syst. Sci. 12 (3), 943–957. https://doi.org/10.5194/hess-12-943-2008.
Thom, A.S., 1975. Momentum, mass and heat exchange of the plant communities. In: Monteith, J.L. (Ed.), Vegetation and Atmosphere. Academic Press, London, pp. 57–109.
Troy, T.J., Wood, E.F., Sheffield, J., 2008. An efficient calibration method for continental-scale land surface modeling. Water Resour. Res. 44 (9). https://doi.org/10.1029/2007wr006513.
Weedon, G.P., et al., 2011. Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. J. Hydrometeorol. 12 (5), 823–848. https://doi.org/10.1175/2011JHM1369.1.
Williams, J.R., Renard, K.G., Dyke, P.T., 1984. EPIC – A new model for assessing erosion's effect on soil productivity. J. Soil. Water. Conserv. 38 (5), 381–383.
Wolff, M.A., et al., 2015. Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study. Hydrol. Earth Syst. Sci. 19 (2), 951–967. https://doi.org/10.5194/hess-19-951-2015.
Yang, X., Magnusson, J., Rizzi, J., Xu, C.Y., 2018. Runoff prediction in ungauged catchments in Norway: comparison of regionalization approaches. Hydro Res. 49 (1). https://doi.org/10.2166/nh.2017.071.
Yang, X.Q., Yong, B., Ren, L.L., Zhang, Y.Q., Long, D., 2017. Multi-scale validation of GLEAM evapotranspiration products over China via China FLUX ET measurements. Int. J. Remote Sens. 38 (20), 5688–5709. https://doi.org/10.1080/01431161.2017.1346400.
Yilmaz, K.K., Gupta, H.V., Wagener, T., 2008. A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrologic model. Water Resour. Res. https://doi.org/10.1029/2007wr006716.