Journal of Hydrology 610 (2022) 127895

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Evaluation of two new-generation global soil databases for macro-scale hydrological modelling in Norway

Shaochun Huang [a,*], Stephanie Eisner [b], Ingjerd Haddeland [a], Zelalem Tadege Mengistu [a]

[a] Norwegian Water Resources and Energy Directorate (NVE), Middelthuns Gate 29, 0368 Oslo, Norway
[b] Norwegian Institute of Bioeconomy Research (NIBIO), Ås, Norway

ARTICLE INFO

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Juan V. Giraldez, Associate Editor

Keywords:
Soil moisture
Norwegian forest soil profile
SoilGrids
Wise30sec
ESA CCI SM
ERA5
GLEAM

ABSTRACT

Lack of national soil property maps limits the studies of soil moisture (SM) dynamics in Norway. One alternative is to apply the global soil data as input for macro-scale hydrological modelling, but the quality of these data is still unknown. The objectives of this study are 1) to evaluate two recent global soil databases (Wise30sec and SoilGrids) in comparison with data from local soil profiles; 2) to evaluate which database supports better model performance in terms of river discharge and SM for three macro-scale catchments in Norway and 3) to suggest criteria for the selection of soil data for models with different complexity. The global soil databases were evaluated in three steps: 1) the global soil data are compared directly with the Norwegian forest soil profiles; 2) the simulated discharge based on the two global soil databases is compared with observations and 3) the simulated SM is compared with three global SM products. Two hydrological models were applied to simulate discharge and SM: the Soil and Water Integrated Model (SWIM) and the Variable Infiltration Capacity (VIC) model. The comparison with data from local soil profiles shows that SoilGrids has smaller mean errors than Wise30sec, especially for upper soil layers, but both soil databases have large root mean squared errors and poor correlations. SWIM generally performs better in terms of discharge using SoilGrids than using Wise30sec and the simulated SM has higher correlations with the SM products. In contrast, the VIC model is less sensitive to soil input data and the simulated SM using Wise30sec is higher correlated with the SM products than using SoilGrids. Based on the results, we conclude that the global soil databases can provide reasonable soil property information at coarse resolutions and large areas. The selection of soil input data should depend on the characteristics of both models and study areas.

1. Introduction

Soil moisture is a key mediator between atmosphere and hydrosphere, which determines the amount of water available for evapotranspiration from land, influencing precipitation and air temperature as feedback (Seneviratne et al., 2010). In addition, it affects streamflow and groundwater recharge via partitioning of precipitation into runoff and infiltration in the hydrological cycle (Orth and Seneviratne, 2013). Despite its importance, soil moisture is rarely studied in Norway, especially for macro-scale catchments. One major obstacle is lack of national-wide soil property information for hydrological or land surface models.

Global soil databases are often considered as alternative soil property information for large-scale modelling and for regions with poor local data. For example, the FAO/UNESCO soil map of the world (FAO, 2003) and the Harmonized world soil database (HWSD) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) are applied as standard inputs in Soil and Water Assessment Tool (SWAT) model (Abbaspour et al., 2019). Dai et al. (2019) reviewed the global soil databases used by the land surface models within the Coupled Model Intercomparison Project 5 (CMIP5). The widely used databases were the FAO/UNESCO soil map of the world, HWSD, the Data and Information System of International Geosphere-Biosphere Programme (IGBP-DIS) database (Global Soil Data Task Group, 2000), the Global Soil Dataset for Earth System Model (GSDE) (Shangguan et al., 2014), etc.

All global soil property databases are generated from soil surveys by one of two general methods (Dai et al., 2019): the linkage method (Batjes, 2003) linking soil profiles and soil mapping units in soil type maps, and the digital soil mapping method (McBratney et al., 2003) using state-of-art machine learning methods to map the spatial distribution of soil properties across the globe. The soil databases produced by

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

https://doi.org/10.1016/j.jhydrol.2022.127895
Received 7 February 2022; Received in revised form 21 April 2022; Accepted 27 April 2022
Available online 4 May 2022

the linkage methods are polygon-based due to the shape of soil type maps while the soil databases derived by digital soil mapping provide gridded, spatially continuous estimates. Thus, the soil databases produced by these methods give distinctively different distribution of soil properties.

Dai et al. (2019) also compared four global soil databases with 94,441 soil profiles from the World Soil Information Service (WoSIS). They found that the SoilGrids system (Hengl et al., 2017), one of the most recently developed databases using the digital soil mapping method, had much better accuracy than HWSD, IGBP-DIS and GSDE, which were developed using the traditional linkage method. Tifafi et al. (2018) compared the soil carbon stock estimates based on the global soil databases with field measurements. They also found that SoilGrids had smaller errors than the World Inventory of Soil Emission Potentials database at a resolution of 30 s (Wise30sec) (Batjes, 2016), which is an improved version of HWSD.

So far, there are limited studies that compared the effects of different soil input data on hydrological or land surface modelling, and evaluations of the new generation of global soil databases are rare. Several studies reported improved model performance in terms of specific hydrological processes, such as soil moisture and river discharge, with better soil information and at higher resolutions (Sheshukov et al, 2011; Guillod et al., 2013; De Lannoy et al., 2014). Livneh et al. (2015) analysed the sensitivities of different soil input data in the mHM model for the Mississippi catchment and concluded that the choice of soil textural properties for a single model can be an appreciable source of uncertainty. In contrast, Ye et al. (2011) and Mukundan et al. (2010) found that the improved high-resolution data did not necessarily improve the streamflow simulation using SWAT model. Hence, the effect of soil input data can be specific for different models, soil data as well as study areas.

All the studies mentioned above mainly compared the simulated river discharge with observations and inter-compared the simulated hydrological processes driven by different soil input data. The simulated soil moisture was rarely validated against observations, probably due to the lack of soil moisture measurements, especially at catchment scales. To fill this gap, satellite-derived soil moisture products have become one important source of surface soil moisture observations in recent years, and they can be used to calibrate hydrological models for ungauged catchments (Choi et al., 2021) or to validate model performance in terms of soil moisture (Kearney & Maino, 2018). In addition, the satellite-derived soil moisture products are often assimilated in land surface or hydrological models to improve the root-zone moisture simulations at large scales (Renzullo et al., 2014; Tian et al., 2019). Based on the advances in observation data, Kearney & Maino (2018) assessed the application of the new generation of global soil data in an infiltration and redistribution model and compared the model outputs with both in-situ soil moisture measurements and satellite-derived soil moisture product. Their results demonstrated the advantages of using the new global soil data for modelling soil moisture at fine spatial and temporal resolution for Australia and encouraged the applications of the new databases for other regions.

In this study, we evaluated two recent global soil property databases (Wise30sec and SoilGrids) generated by different mapping methods for macro-scale hydrological modelling in Norway. To the best of our knowledge, we are the first to compare these two databases with respect to their utility for hydrological modelling, and it is the first study that focuses on macro-scale soil properties and soil moisture simulations in Norway. Different from the previous studies, we applied two hydrological models to account for the uncertainties caused by different model structures, modelling resolutions and descriptions of hydrological processes. In addition, we used a three-step procedure to systematically evaluate the global soil databases. The benchmark data is not only the observed discharge but also the Norwegian forest soil database and three global soil moisture products. Thus, the objectives of this study are 1) to evaluate two recent global soil databases (Wise30sec and SoilGrids) in comparison with data from local soil profiles; 2) to evaluate which database supports better model performance in terms of river discharge and soil moisture for three macro-scale catchments in Norway and 3) to suggest criteria for the selection of soil data for macro-scale hydrological modelling.

2. Study area

Norway is located in Northern Europe and covers an area of about 325,000 km^2. The mainland can be divided into 6 geographic regimes: Finnmark, Nordland, Trøndelag, Western Norway, Eastern Norway and Southern Norway (Fig. 1a). About half of the land area is covered by bedrock and heather (poor shrub vegetations) in high mountains and about 38% of the area is covered by forest.

In this study, we selected three macro-scale catchments with the drainage areas larger than 10,000 km^2 from the Norwegian flood forecasting system (Ruan & Langsholt, 2017), where human activities do not significantly influence the discharge at the outlet. These catchments are the upper Lågen River above the gauging station Losna (ID: 2.145, 11,206 km^2), the upper Glomma River above the gauging station Elverum (ID: 2.604, 15,450 km^2) and the Tana River above the gauging station Polmak (ID: 234.18, 14,170 km^2) (Fig. 1). Both the Losna and Elverum catchments are located in Eastern Norway and the Polmak catchment is partly located in Finnmark and partly in Sweden.

The three catchments span a range of climatic, hydrological and geographic conditions (Table 1). The Losna catchment is located at the highest altitude among the three and it receives the largest annual precipitation amount. About 90% of the precipitation in this catchment ends up as runoff due to the cold climate, steep slopes and large area of bedrock. The Polmak catchment is located in the coldest and northernmost region, and it receives the lowest amount of precipitation per year. It is also flattest and at the lowest altitudes among the three catchments. The Elverum catchment is the largest catchment, and it has the highest annual temperature (above 0 °C) among the three. In both the Polmak and Elverum catchments, about 75% of the precipitation ends up as runoff.

These three catchments represent the dominant land cover types in Norway. In the Losna catchment, heather and bedrock are the dominant land covers (ca. 50%) and forest covers about 25% of the land area. In the Elverum and Polmak catchments, more than 50% of the land area is covered by forest and more than 20% of the land is covered by heather. In addition, due to the large drainage area and distinct locations, these catchments contain more than half of quaternary geological conditions in Norway according to Norway's geological survey.

3. Data and method

In this study, we evaluated Wise30sec and SoilGrids in three steps (Fig. 2). The first step was to compare the global soil data with the Norwegian forest soil database. In the second and third steps, we applied two hydrological models to simulate river discharge and soil moisture using information from the two soil databases separately. The simulated discharge was compared with observed discharge (step 2) and the simulated soil moisture was compared with three global soil moisture products (step 3). In the next subsections, the data and methods are described following the workflow shown in Fig. 2.

3.1. Evaluation 1: Soil characteristics

3.1.1. Global soil property databases

The main characteristics of the two global soil databases are listed in Table 2. Wise30sec (Batjes, 2016) is developed based on the traditional linkage method. It is an improvement of HWSD, as it uses the soil map from HWSD with minor corrections and it incorporates more soil profiles (~21,000 profiles) than HWSD and climate zone maps as categorical covariates. For each soil type, it has seven layers up to 200 cm depth (5 × 20 cm up to 100 cm depth, and 2 × 50 cm up to 200 cm depth) with

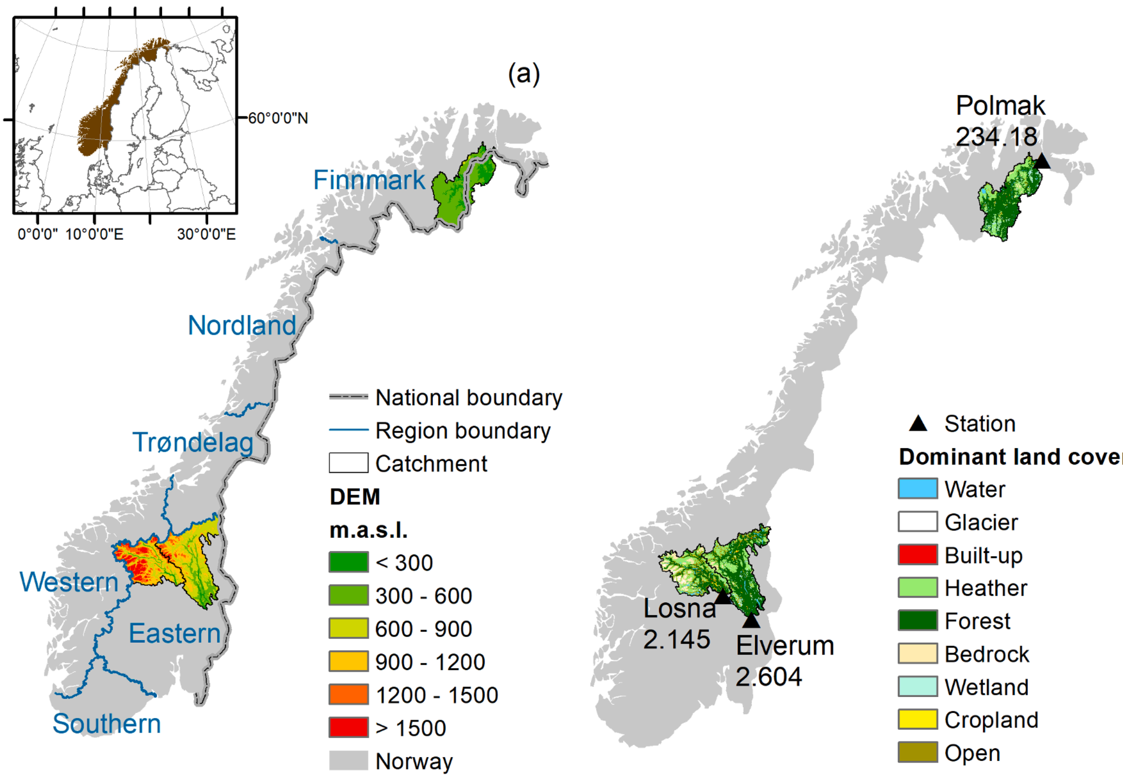

Fig. 1. The digital elevation model (a) and the dominant land cover (b) of the three catchments: Losna (gauge identification (ID): 2.145), Elverum (ID: 2.604) and Polmak (ID: 234.18).

Table 1
The characteristics of the studied catchments.

Gauge ID	Name	Area (km^2)	Elevation [m]	Land cover (%)			Mean temperature [°C]	Mean precipitation [mm/year]	Mean discharge [mm/year]	Runoff coefficient
				Heather	Forest	Bedrock				
2.145	Losna	11,206	1158 (181–2290)	31.0	24.9	18.9	−0.72	810	720	0.89
2.604	Elverum	15,450	819 (182–1915)	22.8	65.2	2.6	0.74	730	539	0.74
234.18	Polmak	14,170	346 (22–952)	25.4	53.7	1.5	−1.55	520	389	0.75

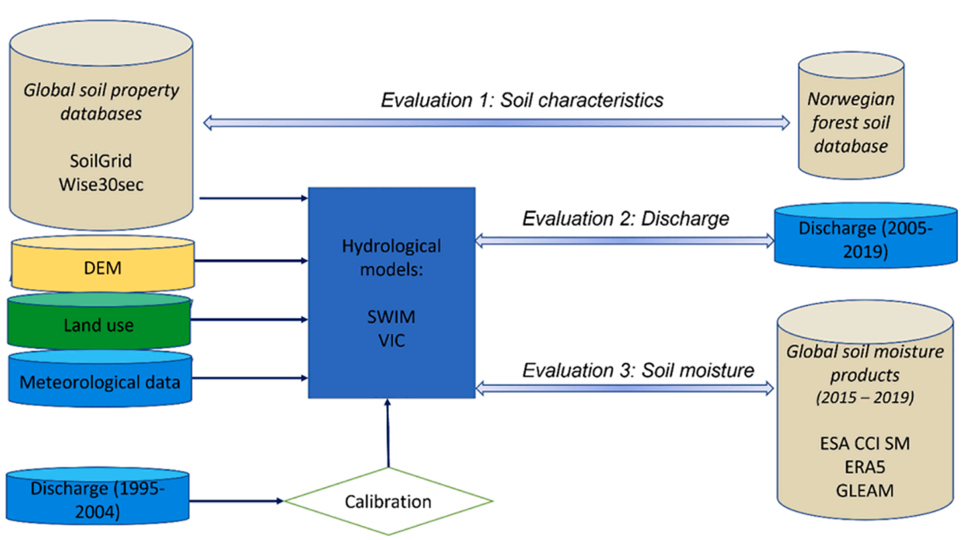

Fig. 2. The flowchart of the three-step evaluation in this study.

Table 2
The main characteristics of Wise30sec and SoilGrids.

Database	Resolution	Number of layers	Number of properties	Depth of the layers (cm)	Mapping method
Wise30sec	1 km	7	20	20, 40, 60, 80, 100, 150, 200	Linkage method
SoilGrids	250 m	6	7	5, 15, 30, 60, 100, 200	Digital soil mapping

uncertainty estimation. The resolution of the data is about 1 km.

SoilGrids (Hengl et al., 2017) is developed using the digital soil mapping method. The latest version released in 2020 was produced by fitting an ensemble of machine-learning methods based on over 230,000 soil profile observations and a series of environmental covariates. Covariates were selected from a pool of over 400 environmental layers from earth observation derived products and other environmental information including climate, land cover and terrain morphology. The outputs of SoilGrids are soil properties at six standard depth intervals (5, 15, 30, 60, 100 and 200 cm) for each grid cell at a resolution of 250 m.

3.1.2. Norwegian forest soil database

The Norwegian forest soil database consists of 1040 sampling sites covering forested land on a 9 by 9 km regular grid, i.e. sampling density follows the geographical distribution of forests in Norway with fewer sites in the northernmost region, central mountain areas, and the coastal areas, where forest is not the dominant land cover (Fig. 3). Profile descriptions, soil sampling and lab analysis for each site were carried out in the period 1988–1992 following standardized procedures (Sveistrup 1984).

In total, 994 soil profiles provide sufficient data to be included in the analysis, i.e. depth and fine earth fractions (sand, silt, clay) for each soil layer. Bulk density for mineral soil layers is estimated using a pedotransfer function from Baritz et al. (2010). The Norwegian forest soil profiles have shallow soil depth in general (Fig. 3a). Only about 20% of the soil profiles are deeper than 80 cm, 50% are deeper than 50 cm and 80% of the soil profiles are deeper than 20 cm. Another major characteristic is that most forest soils belong to sand soils according to the classification by Riley (1996) for Norwegian soils (Fig. 3b).

3.1.3. Comparison between the Norwegian forest soil profiles and global soil data

Based on the characteristics of the soil profile depth, we converted the global soil data and profile data into three layers: layer 1 (0–20 cm), layer 2 (20 – 50 cm) and layer 3 (50–80 cm) to maximize the use of the profile information. As a result, there are about 80%, 50% and 20% of soil profiles available to compare with the global soil data for layer 1, 2 and 3, respectively.

We extracted soil property information from the converted global soil data at the exact locations of the profiles and compared them pairwise. Four key properties (sand, clay, silt content (%) and bulk density (g/cm^3)), which are available for all profiles, are included in the comparison. Three statistic criteria were used: mean error (ME), root mean squared error (RMSE), and coefficient of determination (R2). These statistics were also used in the comparison between global soil data and global soil profiles (Dai et al., 2019).

3.2. Hydrological modelling

In this study, we selected one process-based eco-hydrological model: Soil and Water Integrated Model (SWIM), and one land surface model: the Variable Infiltration Capacity (VIC) model to simulate river discharges and soil moisture. These models were widely used in hydrological modelling for large-scale catchments worldwide and were capable to reproduce river discharges satisfactorily (Huang et al., 2017). In addition, they differ substantially in model structure, process description and spatial resolution of input data, so that we can analyse whether the effects of soil input data differ between the models. Lumped conceptual models were not considered in this study because they often do not require detailed and spatially distributed soil information.

3.2.1. SWIM

SWIM (Krysanova et al., 1998) is an ecohydrological semi-distributed model of intermediate complexity. It was developed based on SWAT-1993 (Arnold et al., 1993) and MATSALU (Krysanova et al., 1989) for meso- and macro-scale catchments. It is driven by six meteorological parameters at daily timesteps: mean, maximum and minimum

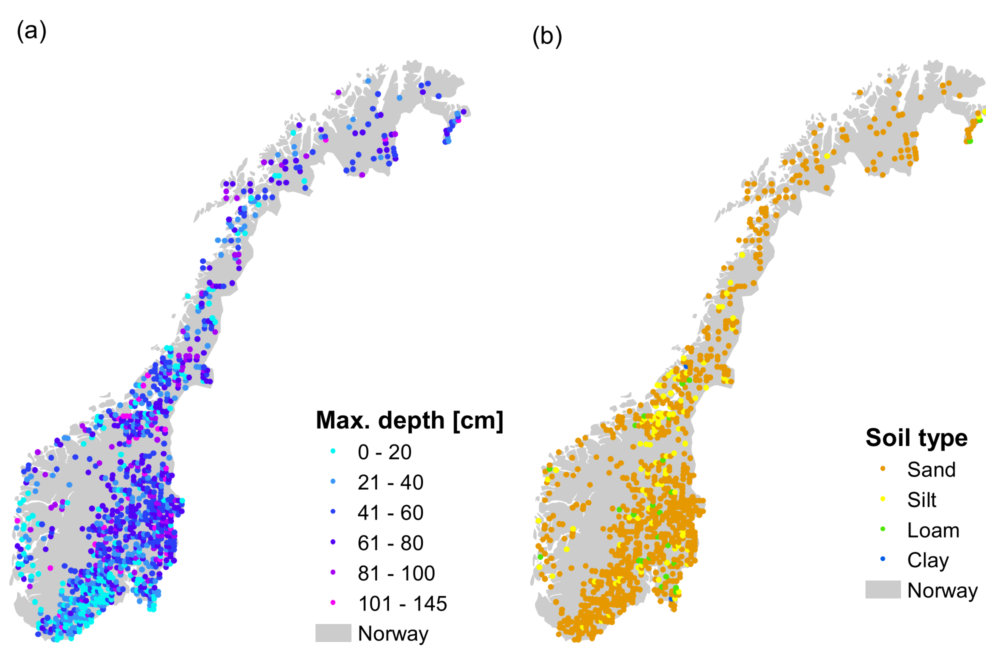

Fig. 3. The maximum depth (unit: cm) (a) and soil types (b) of the Norwegian forest soil profiles.

temperature, precipitation, relative humidity and solar radiation. SWIM has a three-level scheme of spatial disaggregation that is the basin, sub-basins and hydrotopes. The hydrotopes are sets of elementary units in a sub-basin with homogeneous soil and land use types to simulate water flows. In this study, we created regular grid sub-basins at a size of 1 km^2 based on the Universal Transverse Mercator (UTM) zone 33 N coordinate system corresponding to the meteorological forcing data at the same resolution. There are up to five hydrotopes in each sub-basin, with a mean hydrotope area of 0.38 km^2.

Lateral transport of water towards the river network is simulated on the basis of linear storage functions. After reaching the river system, water is routed along the river network to the outlet of the basin using the Muskingum flow routing method (Maidment, 1993). Since the sub-basin is gridded-based and the sub-basin size is small in this study, the Muskingum method can be unstable to rout the runoff from each sub-basin. To overcome this problem, we superimposed the official river network map on the sub-basins and rout the accumulated runoff from a set of sub-basins, which form larger sub-basins based on the river network.

The simulated hydrological system consists of four control volumes: the soil surface, the root zone of soil, the shallow aquifer and the deep aquifer. The soil root zone is subdivided into several layers in accordance with the soil database. The water balance for the soil surface and soil column includes precipitation, surface runoff, evapotranspiration, subsurface runoff and percolation. The water balance for the shallow aquifer includes groundwater recharge, capillary rise to the soil profile, lateral flow and percolation to the deep aquifer.

Surface runoff is estimated as a non-linear function of precipitation and a retention coefficient, which depends on soil water content, land use and soil type (modification of the Soil Conservation Service (SCS) curve number method; Arnold et al., 1990). Lateral subsurface flow is calculated simultaneously with percolation and occurs when the storage in the soil layer exceeds field capacity after percolation.

Potential evapotranspiration (PET) is estimated using the method of Priestley–Taylor (Priestley and Taylor, 1972). Actual evaporation from soil and actual transpiration by plants are calculated separately. An extended degree-day method is used to compute snowmelt (Huang et al., 2013). The percolation from the soil profile to the shallow aquifer is corrected by the delay time function proposed by Sangrey et al. (1984).

3.2.2. VIC

VIC (Liang et al., 1994) (version 4.2.d) is a semi-distributed hydrological model for large-scale applications, solving both the surface energy balance and water balance equations. In this study the model was run in so-called water balance mode, where the energy balance is solved only when snow is present. The land surface processes are modelled at a grid of large cells (0.25° in this study to compare with the global soil moisture products), which can be subdivided into several “tiles” based on land cover types and elevation classes. The fluxes and storages from the tiles are averaged together based on the weights of area fraction to give grid-cell average. The routing of water flow is performed using the unit hydrograph principle. The daily meteorological forcing data used in this study are maximum and minimum temperature, precipitation, longwave radiation, shortwave radiation, atmospheric pressure, vapor pressure and wind speed.

PET is calculated using the Penman-Monteith equation. An energy balance approach is used to represent snow accumulation and ablation on the ground (Andreadis et al., 2009), and the model contains an explicit canopy snow scheme. There are typically 3 soil layers. The top layer is usually 10 cm to ensure reasonable calculation of evaporation from the soil. In this layer, surface runoff and infiltration are controlled by variable infiltration capacity parameterization and soil evaporation is calculated by ARNO formulation (Francini and Pacciani, 1991). The middle layer mainly conveys water to the next layer and the last soil layer generates base flow using the ARNO baseflow formulation (Francini and Pacciani, 1991). All layers within the root zone can lose moisture due to evapotranspiration.

3.2.3. Model input data

The digital elevation model (DEM), soil type and land cover type maps with 1 km horizontal resolution are required to setup the hydrological models in this study. The DEM map is obtained from the Norwegian Mapping Authority and the Swedish National Land Survey.

The soil type maps are required for SWIM and are available from the Wise30sec and SoilGrids databases. The soil type map of Wise30sec shows that there are 72 soil types in Norway while the SoilGrids database estimates 11 most-probable soil types in Norway. Since the soil properties of each most-probable soil type vary between grids, we decided not to use the most-probable soil types from the SoilGrids database but reclassify the soils based on primary soil properties (clay, sand and silt content, organic matter, nitrogen content and bulk density) for all grid cells in Norway using the K-Mean clustering method (Hartigan and Wong, 1979). As a result, we reclassified 70 soil types for the SoilGrids data to have a comparable number of soil types as in Wise30sec. The mean values of soil properties for each new class represent the characteristics of the corresponding soil type.

Besides the primary soil properties available in the soil database, the hydrological models need information on soil parameters such as available water capacity, hydraulic conductivity, etc. We estimated these parameters using the pedotransfer functions by Woesten et al. (1999), which were developed for European soils using clay, silk content, organic matter and bulk density as input. These functions show the overall best performance among the parameter pedotransfer functions for Norwegian soils (Kværno and Haugen, 2011).

Land use information was compiled from the National Land Resource Map (Ahlstrøm et al., 2014) and the remote sensing based forest resource map SAT-SKOG (Gjertsen and Nilsen (2012), supplemented by Corine Land Cover 2000 (https://www.eea.europa.eu/data-and-maps/data/clc-2000-raster-4) for small upstream areas outside the Norwegian borders. 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. The forest types are based on a 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, and are applied here to better reflect spatial variability in hydrologically relevant land surface properties within forested areas. Relevant parameters for each forest type are provided by Majasalmi et al. (2018) (maximum leaf area index in growing seasons and vegetation height) and Bright et al. (2018) (shortwave albedo).

The meteorological forcing data are available at daily time steps and 1 km spatial resolution for all of Norway and areas in neighbouring countries draining to Norway and cover the period 1960 – 2020. Precipitation and temperature (daily mean, maximum and minimum) are obtained from seNorge2018, v20.05 (Lussana et al., 2019), which is an observational gridded dataset. Wind speed is based on the NOrwegian ReAnalysis 10 km (NORA10) product (Reistad et al., 2011). Downward short- and longwave radiation, relative humidity and surface pressure are estimated using the same methods as when preparing the HySN (Erlandsen et al., 2019) and HySN5 (Erlandsen et al., 2021) datasets. HySN is described and compared with surface observations and other data sets in Erlandsen et al. (2019). The dataset prepared for this study, HySN2018v2005ERA5 (https://doi.org/10.5281/zenodo.5947547), is based on Era5 and SeNorge2018, v20.05.

The measured discharge data at the gauges Elverum, Losna and Polmak is used to calibrate and validate the hydrological models. The daily discharge data has been quality checked by the Norwegian Water Resources and Energy Directorate (NVE) and it is continuously available for the period 1995 to 2019.

3.2.4. Model calibration

Ideally, the hydrological models should be calibrated against in-situ soil moisture measurements when the target variable is soil water. However, in-situ soil moisture measurements in Norway are sparse and the data quality is hardly controlled. Hence, we calibrated the hydrological models against discharge at the outlet of the catchments in the period 1995 – 2004. The first year of simulation was used for model spin-up. The dynamically dimensioned search (DDS) algorithm (Tolson and Shoemaker, 2007) was applied to identify the parameterization for each catchment independently. DDS is a stochastic based global search algorithm and designed for calibration problems with many parameters. In this study, we calibrated 11 parameters in the SWIM model: two parameters related to potential evaporation, two related to groundwater flow, two related to soil percolation and saturated conductivity, three related to snow melt and two related to routing. For the VIC model, there are in total nine calibration parameters: four related to soil infiltration and base flow processes, two related to soil depths and three related to snow melt. Detailed information about the calibration parameters is presented in the Appendix.

We selected Nash-Sutcliffe efficiency (NSE) on daily streamflow and its logarithm (LNSE) as the criteria to calibrate the models. These two criteria can show the model performance in terms of both high and low flows, and thus reflect soil functions such as the partitioning of precipitation into surface runoff and infiltration, base flow and the control of water availability for evapotranspiration. Since DDS tries to minimize the objective function (θ), we formulated θ as an equally weighted combination of the differences between the criteria results and their ideal values (1 for NSE and LNSE) (Eq. (1)).

$$\theta = (1 - NSE) + (1 - LNSE) \tag{1}$$

For each catchment, the hydrological models were calibrated twice using the same meteorological forcing data, the same calibration parameters and their ranges, and the same objective function. The only difference is the prescribed soil input data, which is based on either SoilGrids or Wise30sec data.

3.3. Evaluation 2: Discharge

Evaluation of discharge is mainly based on the statistic criteria NSE, LNSE and the percent bias of water balance (PBIAS) on daily and monthly streamflow at the outlet gauges in the validation period (2005 – 2019). In addition, visual inspection of monthly hydrographs helps to evaluate model performance in terms of seasonal dynamics.

3.4. Evaluation 3: Soil moisture

Three global soil moisture products were compared to simulated soil moisture from the hydrological models. They include one solely satellite-based product: the European Space Agency Climate Change Initiative Plus Soil Moisture (ESA CCI SM) (Gruber et al., 2019) and two products from models with satellite data assimilation: ERA5 (Hersbach et al., 2020) and the Global Land Evaporation Amsterdam Model (GLEAM, version 3.5b) (Martens et al., 2017). These products are publicly available and show comparable performance compared with more than 800 in-situ soil moisture measurements worldwide (Beck et al., 2021).

3.4.1. ESA CCI SM product

The most recent ESA CCI SM product is the ESA CCI SM v06.1 published in 2021 and it is the most accurate ESA CCI SM product at present. The combined ESA CCI SM product is a merged data from various satellites, including three active and ten passive satellites. It provides long-term global data from 1978 to 2020 at a regular grid of 0.25°. In this study, we only applied the data after 2015 because one of the newest satellites, the Soil Moisture Active Passive (SMAP) satellite, provides data from 2015.

Since the soil moisture measured by satellites only represents the condition of top 2–3 cm soil layer, we calculated the Soil Wetness Index (SWI, Eq. (2)) on the satellite time series using the exponential smoothing filter (Wagner et al., 1999) to compare with the hydrological model outputs for the top 10 cm soil layer.

$$SWI(t_n) = \frac{\sum_i^n SM_{sat}(t_i) e^{-\frac{t_n - t_i}{T}}}{\sum_i^n e^{-\frac{t_n - t_i}{T}}} \tag{2}$$

where SMsat (m^3/m^3) is the soil moisture retrieval at day t_i, T represents the time lag constant and was set to 5 days. Following Pellarin et al. (2006), the SWI at time t_n was only calculated if ≥ 1 retrievals were available in the interval $[t_n - T, t_n]$ and ≥ 3 retrievals were available in the interval $[t_n - 3T, t_n - T]$.

3.4.2. ERA5

ERA5 is the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis. Compared to previous generation reanalysis, ERA5 provides an enhanced number of output parameters at spatial resolution of 0.25° and hourly temporal resolution from 1979 to present. ERA5 includes an advanced land data assimilation system to analyse land surface prognostic variables. It applies the soil texture data from the digital soil map of the world and derived soil properties (FAO, 2003) and ASCAT soil moisture data for data assimilation. The output of volume of water in soil layer 1 (0–7 cm) is used for comparison in this study.

3.4.3. GLEAM

GLEAM is a set of algorithms estimating terrestrial evaporation and root-zone soil moisture from satellite data from 1980 to 2020 at 0.25° resolution. It uses the ESA CCI SM (v5.3) for data assimilation. The soil properties used in GLEAM come from the database of Global Gridded Surfaces of Selected Soil Characteristics generated by IGBP-DIS (Global Soil Data Task Group, 2000). The surface layer of GLEAM (0–10 cm) is used for comparison.

3.4.4. Evaluation approach

Since satellite soil moisture products often do not provide any information when soil is frozen or snow-covered, we compared the data only in July and August, when the ground is snow-free in all catchments. Similar to previous studies (e.g., Beck et al., 2021), we used the Pearson correlation coefficient (r) to calculate the correlations between the original soil moisture products and the simulated soil moisture. For the satellite product, we only calculated the correlations if more than 200 soil moisture estimates were available. The simulated soil moisture by SWIM was aggregated to the grid resolution of the VIC model (0.25°) and the soil moisture products to make the comparison consistent. In addition, we calculated r for the low- and high-frequency fluctuations of the soil moisture time series as suggested by Gruber et al., (2020). The low-frequency fluctuations are the average soil moisture time series using a 30-day central moving window. The moving average was calculated when more than 21 days with values were available in the 30-day window. The high-frequency fluctuations were calculated by subtracting the low-frequency fluctuations from the original time series.

4. Results

4.1. Evaluation 1: Soil characteristics

Table 3 shows all statistical results for three soil layers between the Norwegian forest soil profiles and the extracted soil properties at the profile locations from SoilGrids and Wise30sec. In general, both soil databases underestimate sand content and overestimate clay content in all layers. SoilGrids has better accuracy than Wise30sec for layer 1 and layer 2 shown by ME and RMSE, but Wise30sec shows smaller ME and RMSE for layer 3 in terms of sand and silt content.

The MEs are less than 10% for all particle-size properties in both soil databases. The MEs of sand, silt and clay content in SoilGrids are less

Table 3
The mean error (ME), root mean squared error (RMSE) and coefficient of determination (R2) of the two global soil data compared with the Norwegian forest soil profiles.

		Layer 1 (0–20 cm)			Layer 2 (20–50 cm)			Layer 3 (50–80 cm)		
		ME	RMSE	R2	ME	RMSE	R2	ME	RMSE	R2
Bulk density (g/cm^3)	SoilGrids	−0.09	0.50	0.00	0.03	0.51	0.00	0.14	0.63	0.01
	Wise30sec	0.13	0.51	0.00	0.14	0.52	0.00	0.30	0.66	0.00
Sand (%)	SoilGrids	−3.88	15.53	0.07	−4.74	16.36	0.05	−9.65	18.47	0.00
	Wise30sec	−8.45	17.54	0.01	−5.39	16.92	0.02	−7.57	17.37	0.01
Silt (%)	SoilGrids	0.47	13.38	0.08	0.19	13.99	0.04	4.26	13.97	0.00
	Wise30sec	2.15	14.13	0.01	−0.40	14.35	0.01	2.03	13.93	0.00
Clay (%)	SoilGrids	3.28	5.76	0.02	4.47	6.52	0.06	5.23	7.01	0.01
	Wise30sec	6.18	7.48	0.00	5.74	7.37	0.00	5.46	6.82	0.00

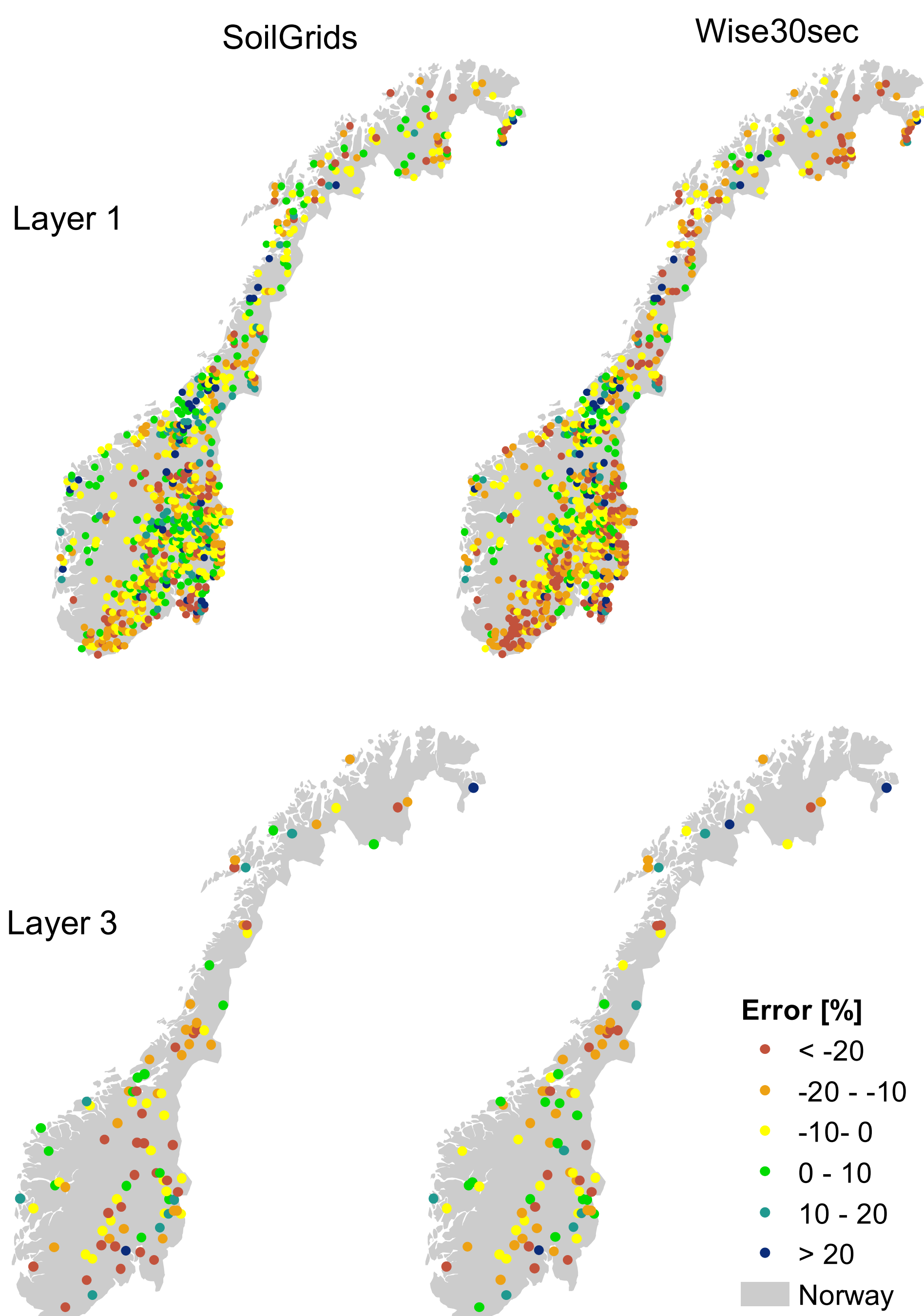

Fig. 4. The error of sand content for soil layer 1 (0–20 cm; upper panels) and layer 3 (50–80 cm; lower panels) in the SoilGrids (left panels) and the Wise30sec (right panels) databases compared with the Norwegian forest soil profiles.

than 5% in the upper two layers and the MEs of silt content are always less than 5% for all layers in both databases. However, the RMSEs of sand and silt content are larger than 10% in both soil databases for all layers, indicating considerable errors at individual points. In addition, Table 3 shows that there are poor correlations (R2 less than 0.1) between the global soil data and the soil profiles for all studied soil properties.

The high RMSEs and poor correlations are attributed to poor representation of both spatial and magnitude variation in the global soil data. Fig. 4 shows the errors of sand content in layer 1 and layer 3 as an example. The random spatial distributions of the errors indicate no clear error pattern in both soil databases and layers because the measured soil properties can differ substantially between adjacent profiles and the global soil data cannot capture such high spatial variability. In addition, Fig. 5 shows that the measured soil properties are much more

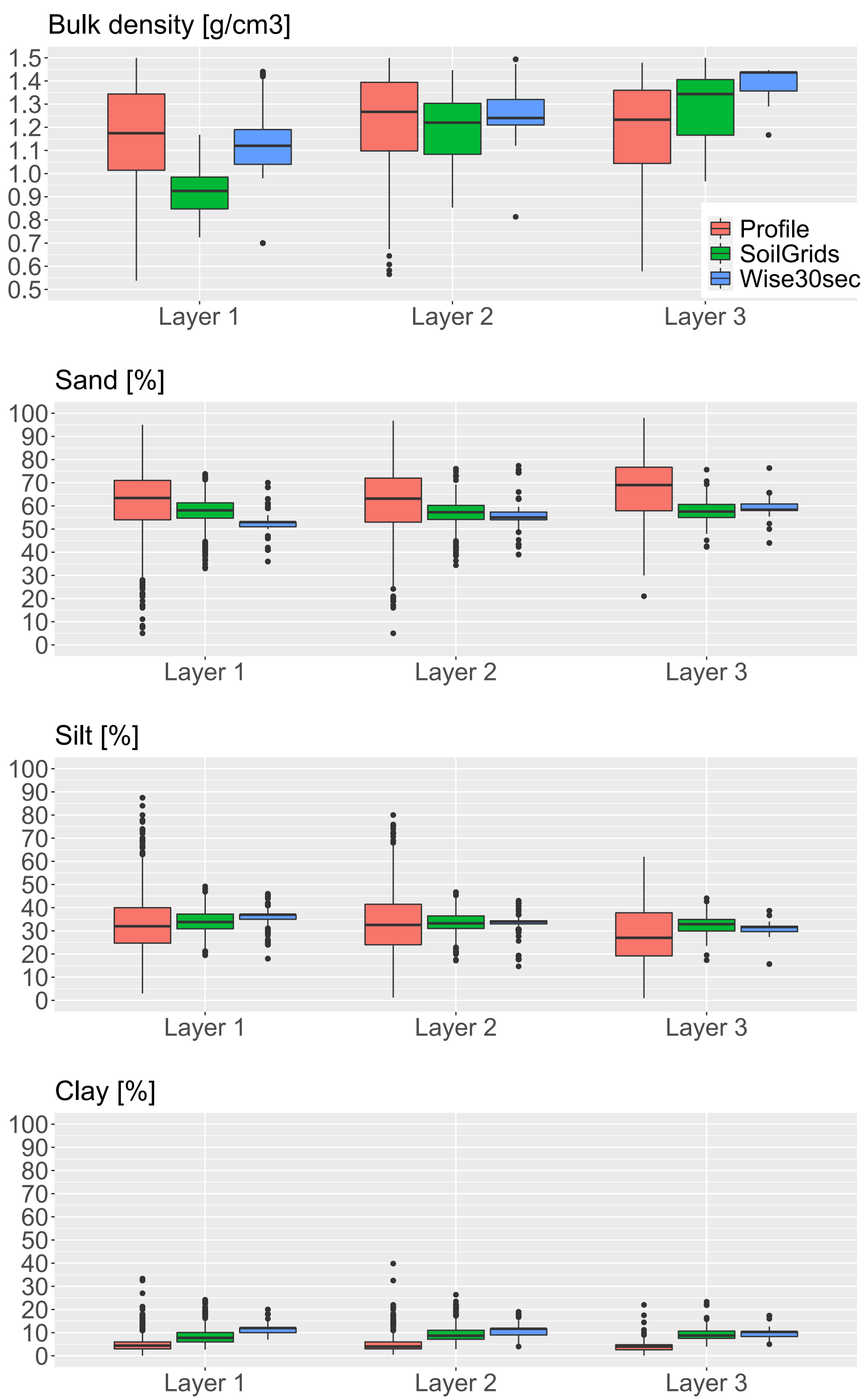

Fig. 5. The boxplots of bulk density, sand, silt and clay content for three soil layers in the Norwegian forest soil profiles and the extracted SoilGrids and Wise30sec data at the same profile locations.

heterogeneous than the estimates from the global soil databases. For example, the measured sand content ranges from 5 to 100%, while the estimates from the global soil databases spread mainly between 40% and 80%.

Bulk density, which represents the soil morphological properties and is an indicator of soil porosity, also shows discrepancies between the global data and the estimates for the forest soil profiles. Both soil databases overestimate the bulk density for all layers, except SoilGrids for layer 1 (Table 3). The largest overestimation is found for layer 3, with ME of 0.14 and 0.3 g/cm^3 for SoilGrids and Wise30sec, respectively. The RMSE for both soil databases and all layers are larger than 0.5 g/cm^3, which equals to about 19% of the particle density (normally assumed to be 2.65 g/cm^3).

The high RMSE and poor correlation results indicate that both global databases have significant biases compared with individual profiles. Hence, the estimated soil water properties based on the global databases can be uncertain for hydrological modelling at small scales. However, the small ME results indicate that the global soil databases may still be capable to provide reasonable aggregated estimates for large grid cells or large catchments, especially for topsoil layers. Therefore, evaluations 2 and 3 only focus on the results at catchment and large grid scales in the following sections.

4.2. Evaluation of model calibration results

The SWIM and VIC models were set up for each catchment separately. For SWIM, two soil property data were generated based on the global soil databases. In addition, there are two hydrotope input files for each catchment, which include different soil type information corresponding to the soil property data. For VIC, two soil input files were prepared for each catchment. Since SoilGrids outperforms Wise30sec mainly for topsoil, we assumed that both databases had similar reliability as input for macro-scale hydrological modelling and calibrated each model for each catchment using different soil input files separately. The aim of the comparison is to investigate which soil database supports a better model performance in terms of discharge.

Table 4 lists the calibration results in terms of NSE, LNSE and PBIAS. In general, both models can reproduce the daily discharge well with NSE/LNSE larger than 0.74 and absolute PBIAS less than 10%, except VIC using Wise30sec for Polmak catchment. These results confirm that the global soil databases can be a good alternative for use in macro-scale hydrological modelling in Norway.

4.3. Evaluation 2: Simulated discharge

We mainly compared the effects of soil data input in the validation period (Table 4). For SWIM, the absolute PBIAS are good (less than5%) for all catchments with both soil databases, but the NSE and LNSE show different model performance depending on the characteristics of catchments and the choice of soil databases. SWIM performs well with daily NSE/LNSE greater than 0.8 for the Losna catchment regardless which soil database is applied, probably due to the large portion of bedrock in this catchment. For the other two catchments, SWIM performs generally better with SoilGrids data than with Wise30sec. The daily NSE and LNSE using SoilGrids are larger than the ones using Wise30sec by 0.05 – 0.09 except the NSE for Polmak, while the monthly NSE and LNSE using SoilGrids are larger than the ones using Wise30sec by up to 0.08.

Different from the SWIM results, the simulated discharge by VIC shows no significant difference in both daily and monthly NSE and LNSE ($\leq$0.03) between the two soil input databases for the Losna and Elverum catchments. However, there are large negative PBIAS (down to −13%) for these catchments. For the Polmak catchment, VIC has better model performance using SoilGrids than using Wise30sec, especially for low flows.

Fig. 6 illustrates the comparison between observed and simulated long-term mean monthly discharge in the validation period. Similar to the results above, there is almost no difference between simulations driven by the two soil databases for SWIM on long-term mean monthly discharges at Losna. For the other two catchments, there are mainly differences in the peaks in May. VIC reproduces similar hydrographs with both soil data for all three catchments, except in late summer for Polmak.

4.4. Evaluation 3: Soil moisture

The Pearson correlation coefficient (r) was calculated for each grid cell between the simulated outputs and the soil moisture products from 2015 to 2019. Fig. 7 summarizes the results of r for all grid cells, which have more than 80% of the cell area within the catchment boundaries. It shows that both the SWIM and VIC outputs have much better correlation with the ERA5 and GLEAM products than the ESA CCI SM data, and the SWIM outputs have better agreements with the ERA5 products than the VIC ones. Since the objective of this study is to evaluate the differences between the global soil databases rather than between the soil moisture products or hydrological models, we only focus on the difference of r between SoilGrids and Wise30sec in the following results.

For SWIM, the soil moisture outputs using SoilGrids have generally higher correlations with all original soil moisture products and their low-frequency fluctuations than the outputs using Wise30sec. The differences of median r for the original data are less than 0.03 while the differences of median r for the low-frequency fluctuations range from 0.03 to 0.09 between SoilGrids and Wise30sec. In contrast, the median r for the high-frequency fluctuations using Wise30sec is higher than the

Table 4
The statistic criteria of the SWIM and VIC model performance using the SoilGrids and Wise30sec data as input for the three catchments in both calibration (1995 – 2004) and validation (2005 – 2019) periods. NSE: the Nash-Sutcliffe efficiency on daily/monthly streamflow, LNSE: the logarithm Nash-Sutcliffe efficiency on daily/monthly streamflow and PBIAS: the percent bias of water balance.

Hydrological model	Soil data	Gauge	Calibration (1995–2004)					Validation (2005–2019)				
			NSE (daily)	LNSE (daily)	NSE (monthly)	LNSE (monthly)	PBIAS	NSE (daily)	LNSE (daily)	NSE (monthly)	LNSE (monthly)	PBIAS
SWIM	SoilGrids	Elverum	0.84	0.82	0.92	0.86	1	0.81	0.77	0.88	0.83	−4
		Losna	0.86	0.80	0.92	0.83	−1	0.83	0.84	0.88	0.88	−4
		Polmak	0.88	0.88	0.95	0.93	−2	0.65	0.76	0.89	0.85	2
	Wise30sec	Elverum	0.77	0.75	0.89	0.82	0	0.74	0.68	0.86	0.76	−5
		Losna	0.86	0.79	0.92	0.82	0	0.84	0.83	0.90	0.87	−3
		Polmak	0.78	0.85	0.94	0.92	−2	0.66	0.71	0.81	0.82	2
VIC	SoilGrids	Elverum	0.80	0.84	0.90	0.89	−8	0.70	0.78	0.83	0.85	−13
		Losna	0.84	0.87	0.93	0.92	−7	0.77	0.86	0.88	0.91	−10
		Polmak	0.74	0.85	0.92	0.90	−5	0.70	0.67	0.86	0.76	2
	Wise30sec	Elverum	0.82	0.84	0.91	0.90	−7	0.73	0.77	0.85	0.82	−12
		Losna	0.81	0.84	0.90	0.90	−7	0.76	0.84	0.86	0.89	−10
		Polmak	0.72	0.64	0.86	0.71	2	0.64	0.32	0.75	0.43	9

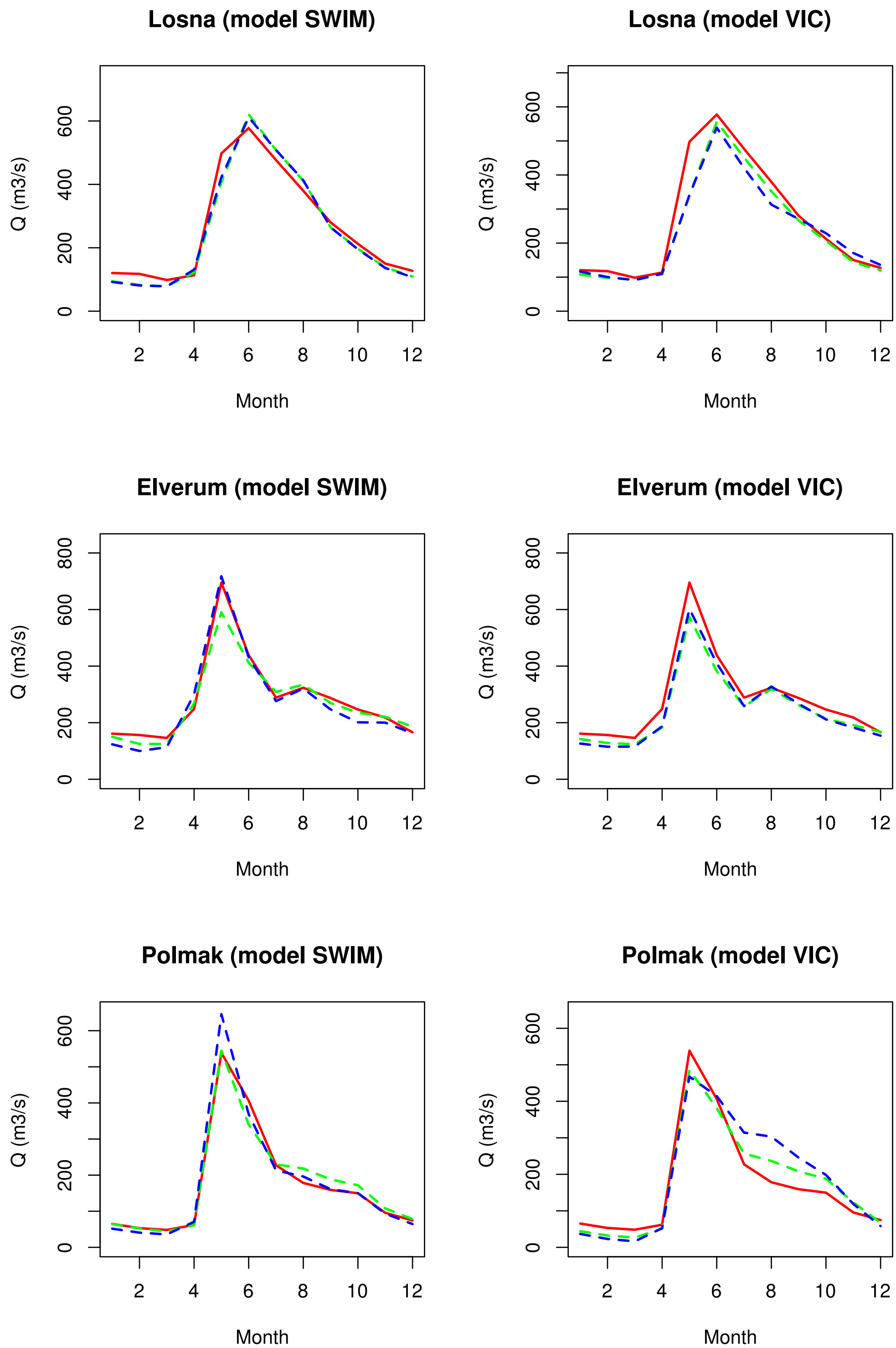

Fig. 6. Comparison between the simulated and observed long-term mean monthly discharge in the validation period (2005 – 2019) using the SWIM and VIC models. The red solid lines represent observed discharge, and the green and blue dashed lines represent the simulated discharge using SoilGrids and Wise30sec data, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

median r using SoilGrids by up to 0.04 when compared with ERA5 and GLEAM. However, the correlations for the high-frequency fluctuations between the simulated and ERA5/GLEAM data are generally poorer than the correlations for the original data and low-frequency fluctuations.

For VIC, the soil moisture outputs using Wise30sec have higher correlations with all global soil moisture products than the outputs using SoilGrids except for the GLEAM high-frequency fluctuations. The differences of median r between the two soil databases are most pronounced when the simulated results are compared with the ESA CCI SM data, up to 0.05. The differences of median r between the soil databases are less than 0.03 when compared with the ERA5 and GLEAM outputs.

Figs. 8 and 9 illustrate the results of r between the SWIM/VIC outputs and the original soil moisture products for each grid cell to identify the regions with poor model performance. Interestingly, there are no distinct spatial distribution of r between SoilGrids and Wise30sec except

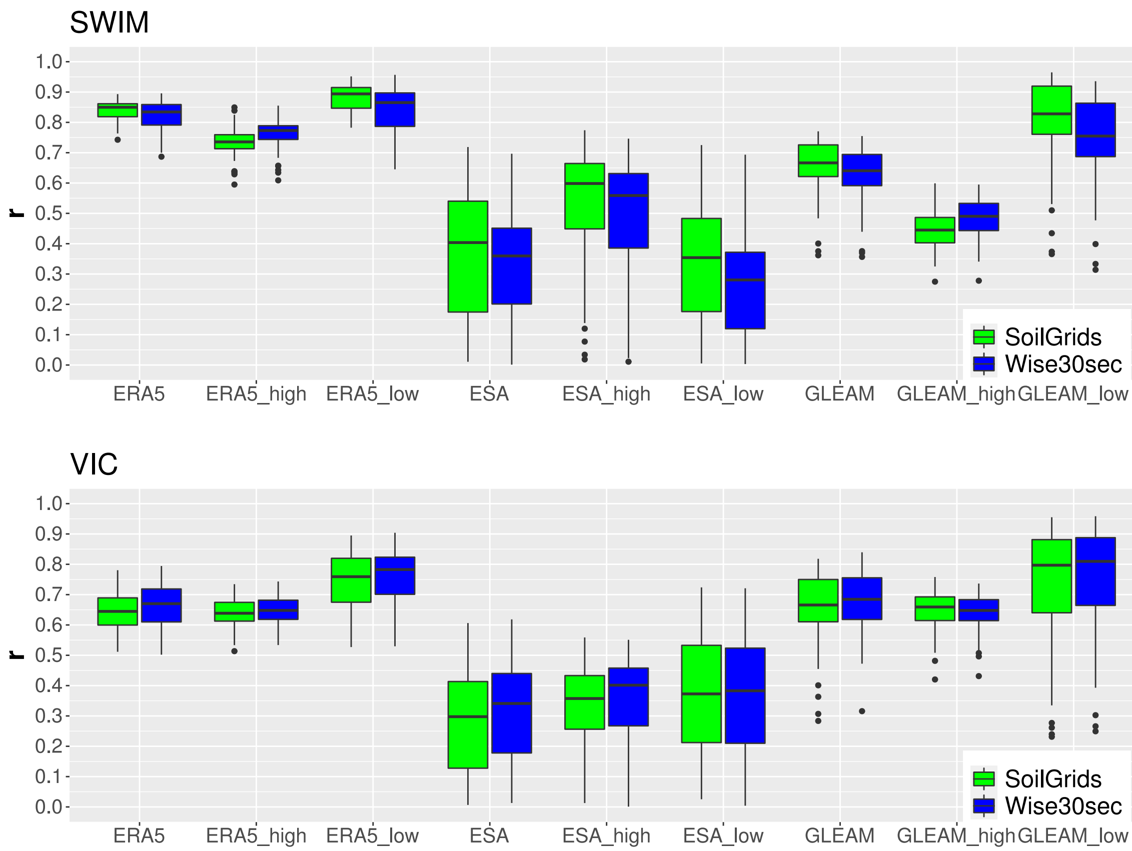

Fig. 7. The Pearson correlation coefficient (r) between the SWIM soil moisture outputs and three global soil moisture products (ERA5, ESA CCI (indicated by ESA in the figure) and GLEAM) (upper panel) and between the VIC soil moisture outputs and the same soil moisture products (lower panel) for all grid cells, which have more than 80% of cell areas within the catchment boundaries. "_high" and "_low" represent the high- and low- frequency fluctuations of the original soil moisture time series (see Section 3.4.4 for their definitions), respectively.

in the upper part of the Polmak catchment for SWIM outputs. In Fig. 8, the SWIM outputs show good correlation (r greater than 0.7) with the ERA5 data for all grids. Poor correlations (r less than 0.3) are mainly found in most part of the Elverum catchment, the upper part of the Losna catchment and the lower part of the Polmak catchment between the simulated soil moisture and the ESA CCI SM data. The SWIM soil moisture data does not agree well with the GLEAM data (r less than 0.5) in the upper part of the Losna catchment.

The VIC outputs have different correlations with the global soil moisture products in terms of both spatial distribution and magnitude compared with the SWIM outputs (Fig. 9). The outputs using Wise30sec have higher correlations with ERA5 than the ones using SoilGrids in the lower part of Elverum and the upper part of Polmak. The correlations with the ESA CCI data are lower than 0.3 in most grids in Eastern Norway and the lower part of Polmak, especially using SoilGrids as input. The correlations with the GLEAM data are generally good in the Polmak catchment using both soil input data but are poor in the upper part of the Losna catchment using the SoilGrids data.

5. Discussion

5.1. Data quality

This study used a three-step evaluation on two global soil databases and their effects on discharge and soil moisture simulations. Such multiple-step evaluation gives valuable information on soil data quality and serves as a basis for selecting input data for macro-scale hydrological or land surface modelling in general. The approach chosen in this study corresponds to the enhanced calibration/validation strategies for hydrological modelling suggested by Krysanova et al. (2018), i.e., to evaluate the quality of input data and consider the uncertainty of input data before model calibration. However, it should be acknowledged that the benchmark data used in this study also contain several weaknesses and uncertainties.

One of the important benchmark datasets in this study is the Norwegian forest soil database. The profiles provide a good spatial distribution of soil properties because forest is one of the dominant land covers in Norway. However, the forest soil profiles cannot represent the soil conditions for other land use types. For example, the European HYdropedological Data Inventory (EU-HYDI) (Weynants et al., 2013) shows that there is more clay and silt content in Norwegian agricultural soils than in forest soils and more than half of the agricultural soil profiles are deeper than 1 m. The disadvantage of the EU-HYDI is lack of sufficient information on the Norwegian profile locations and thus it cannot be applied for pairwise comparison as done in this study.

The global soil moisture products are another benchmark data in this study. Compared with the in-situ soil moisture measurements, they do not perform adequately for all climate, vegetation cover and topographic conditions. Beck et al. (2021) showed that all three products used in this study have weaker performance in cold climate than in temperate climate. Another global assessment of various satellite surface soil moisture products has also shown low correlations (median r ranges from 0.4 to 0.6) with in-situ soil moisture measurements in subarctic climate (Ma et al., 2019). In addition, it is challenging to correctly derive soil moisture data from satellites in dense vegetation areas and steep terrain (Beck et al., 2021, Blyverket et al., 2019). These problems increase the uncertainty of satellite products for Norway, especially in forested and mountainous regions. Given the large uncertainties of the global soil moisture products, we did not focus on the individual model performance in terms of soil moisture but aimed to compare the effects

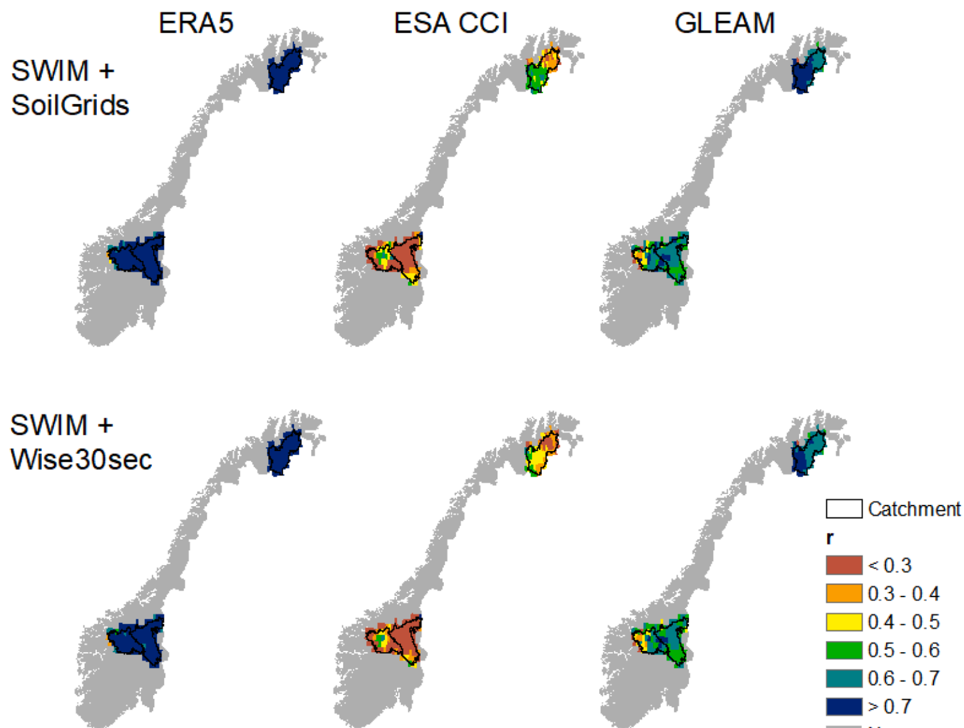

Fig. 8. The Pearson correlation coefficient (r) for each grid cell between the SWIM soil moisture outputs and the original global soil moisture products ERA5 (left panels), ESA CCI (middle panels) and GLEAM (right panels). The upper panel shows SWIM outputs with SoilGrids and the lower panel shows SWIM outputs with Wise30sec.

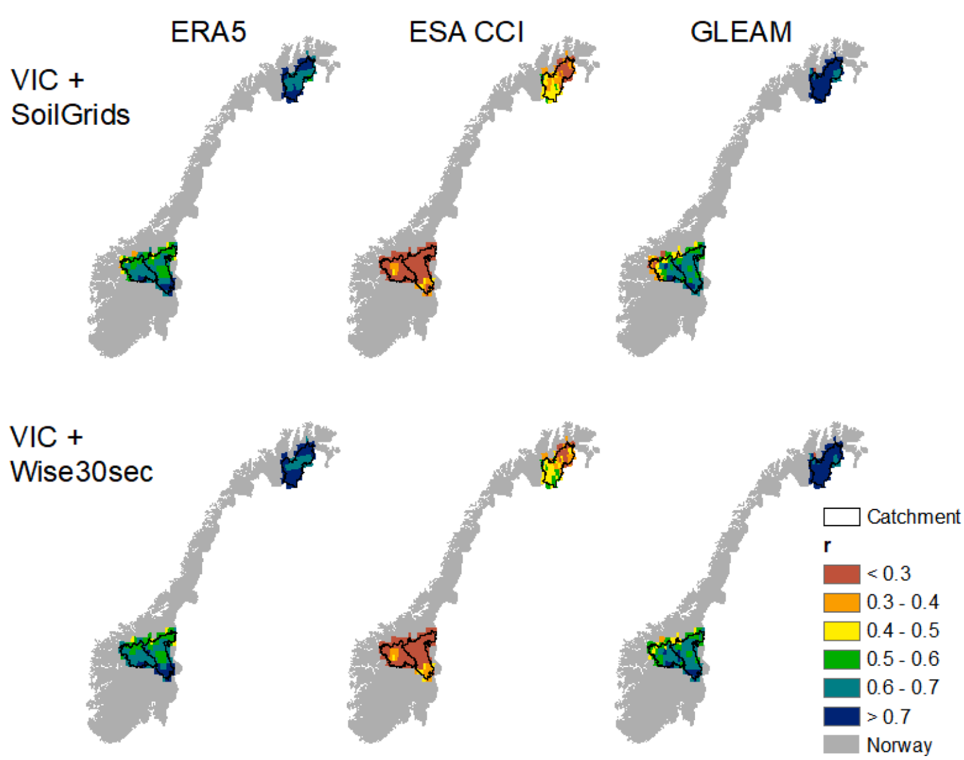

Fig. 9. Te same as Fig. 8 but the Pearson correlation coefficient (r) between the VIC model outputs and the global soil moisture products.

of different soil property data on hydrological modelling.

The results of this study show that the most updated global soil data still contain considerable discrepancies at point or small scales in Norway in comparison with data from local soil profiles. One probable reason is that the global soil data were derived based on a global database of soil profiles and the percentage of Norwegian soil profiles is minor. For example, Norwegian soil profiles account for only 0.2% in the most recent world soil profile database (Batjes et al., 2020). Hence, the estimated soil properties in Norway can be influenced by soil profiles from other parts of the world, especially using the linkage method. In addition, there are fewer data for soil physical than soil chemical attributes, and fewer measurements for deeper than for surficial horizons in most of the soil profile data worldwide (Batjes et al., 2020). The estimation methods also play an important role for the underestimation of spatial variation in soil properties in the global databases. The soil properties in SoilGrids vary smoothly within small areas due to the grid-based spatially continuous estimation while the soil properties in Wise30sec are constant within each soil type polygon.

Soil depth can also be a source of uncertainty in the SWIM simulations. The soil depth in the global databases is always 2 m, but we only applied the data in the top 1 m soil because of the shallow soil depths indicated by the Norwegian forest soil profiles. The prescribed soil depth of 1 m should be appropriate to simulate the three catchments in this study because the calibrated soil depths in VIC range from 0.9 to 1.2 m. However, this prescribed soil depth may not be appropriate for other catchments, and it can be improved in the future with variable soil depths based on the bedrock depth information (Pelletier et al., 2016).

Finally, we should keep in mind that lack of national soil property maps and in-situ soil moisture measurements are major obstacles for soil moisture studies in Norway. We failed to bias correct the global soil data based on the forest profiles because the profiles contain much larger heterogeneity in space and magnitude than the global data. In addition, a soil moisture network and quality-controlled data are expected to improve the understanding of soil water processes in Norway.

5.2. Uncertainty of hydrological modelling

In this study, SWIM using SoilGrids shows a better performance in terms of river discharge and soil moisture than using Wise30sec for most cases. VIC using Wise30sec tends to give better soil moisture estimations than using SoilGrids, but the effects of soil input data on VIC are not as strong as on SWIM. The different effects of soil input data on the hydrological models can be attributed to the prescribed vs calibrated soil parameters, the land surface processes, model resolution and catchment characteristics.

Firstly, there is less freedom for SWIM to change soil water processes than for VIC during calibration. For SWIM, there are six-layer prescribed soil information, including depth, clay, sand silt, bulk density, porosity, available water content, field capacity and saturated conductivity. Among them, only the saturated conductivity can be corrected globally during calibration. In contrast, there are only three soil layers in VIC, including the prescribed information such as bulk density, saturated conductivity and soil moisture content at critical and wilting points. There are four calibration parameters to adjust infiltration and base flow simulations. Due the different number of prescribed and calibrated parameters, the prescribed soil data play a bigger role on SWIM than on VIC for soil water.

Secondly, the simulated land surface fluxes, such as evaporation and snow melt, can dominate the simulation of soil moisture and discharge compared with soil parameters in some cases. For example, snow scheme/parameters are probably more important than soil scheme/parameters during the snow melt season because they directly control the amount of water above ground. The degree day factor snow scheme in SWIM is calibrated to match snowmelt, whereas the energy balance snow scheme in VIC is not calibrated and hence it is more difficult to match snowmelt that infiltrates into soil.

Thirdly, the loss of topographic variability in the VIC due to grid cell size (0.25°) can also contribute to weaker model performance in terms of discharge. VIC accounts for subgrid-scale variability in atmospheric forcings, which reduces its sensitivity to spatial scale compared to using cell averaged forcings at coarse spatial resolutions (Haddeland et al., 2002; Boone et al., 2004). However, 0.25° spatial resolution is most likely too coarse to capture all relevant processes in these catchments. Magnusson et al. (2019) found that there is an increasing error with spatial resolution for snow water equivalent simulations in Norway using the Flexible Snow Model, especially between 1 km and other coarse resolutions. In addition, the aggregated soil information at large grid cells may also reduce the differences between the two global soil databases and thus diminish the effects of soil input data on hydrological modelling. In contrast, the sub-basin size for SWIM is only 1 km^2 so that the model can use the input data directly, without a loss in spatial variability of the meteorological, topographic and soil information.

Fourthly, the effects of soil input data are also dependent on the characteristics of the study areas. For bedrock dominant regions such as the Losna catchment, the soil information is less important than for other catchments irrespective of the hydrological model used because a large share of rainfall and snowmelt will form surface runoff on impermeable surfaces. Hence, the criteria for selecting soil input data should consider the characteristics of both models and study areas. For the more physically based soil modules such as in SWIM, it is highly recommended to evaluate the quality of the soil data to reduce the input data uncertainty.

Lastly, we should acknowledge that the calibration against only river discharge at the outlet of catchments may lead to equifinality problems for the calibrated parameters. However, the aim of this study is not to provide the best simulations for the three catchments, but to compare the effects of different global soil data as input. Such global calibration at a single point allows us to detect the potential problems of the spatial distribution in the soil input data, while calibration at several intermediate gauges might hide the problems of input data by tuning the calibration parameters for different regions separately.

6. Summary and conclusions

In this study, we evaluated two state-of-the-art global soil databases (SoilGrids and Wise30sec) and their effects on macro-scale hydrological modelling in Norway in three steps. Firstly, we compared soil texture and bulk density between the Norwegian forest soil profiles and the global data at the same locations. The results show that SoilGrids has smaller MEs than Wise30sec for all soil properties in the upper soil layers and for bulk density and clay content in the deep soil layer. However, both soil databases have large RMSE (greater than 10% for sand and silt content and greater than 0.5 g/cm^3 for bulk density) and poor correlations (R2 less than 0.1), mainly because they cannot capture the high spatial and magnitude variability in physical soil properties found in the local soil profiles.

Secondly, we evaluated the effects of different soil input data on river discharge simulations for three macro-scale catchments in Norway using one process-based hydrological model SWIM and one land surface model VIC. SWIM generally performs better using the SoilGrids data as input than using the Wise30sec data, with higher NSE and LNSE by up to 0.09 in most cases. The VIC outputs are not sensitive to soil input data for the two catchments in Eastern Norway with marginal differences in NSE and LNSE ($\leq$0.03) between the simulations based on two soil databases. For the snowmelt dominated Polmak catchment, VIC performs better using SoilGrids than using Wise30sec, especially for low flows.

Thirdly, we compared simulated soil moisture with three global soil moisture products (one from a merged satellite product (ESA CCI SM v6.1) and two from global models with data assimilation (ERA5 and GLEAM)) at grid cells of 0.25°. For SWIM, the soil moisture outputs based on SoilGrids have generally higher correlations with original soil moisture data and their low-frequency fluctuations than the outputs based on Wise30sec. The differences in median r are up to 0.09 between the simulations based on two soil databases. In contrast, VIC simulated soil moisture using Wise30sec has higher correlations with the ERA5 and GLEAM data than using SoilGrids by up to 0.05 in terms of median r.

Based on the three-step evaluation, we can conclude that 1) the global soil databases are capable to provide reasonable soil property information at coarse resolutions for macro-scale hydrological modelling in Norway, but they can introduce considerable errors at point or small scales; 2) SoilGrids contains smaller mean errors when compared with the Norwegian forest soil database and supports better model performance than Wise30sec in terms of discharge and soil moisture in most cases simulated by SWIM; 3) both global soil databases can be used for VIC as input and the selection should be based on the study of interest and the characteristics of study area.

The results of this study can serve as a basis for other macro-scale hydrological or land surface models to select the soil input data for Norwegian studies. The multi-step evaluation can also be applied for other geographic regions and is highly recommended for soil process-based models, such as SWIM, in the applications with poor local soil information. In addition, this study evokes the need of national soil property maps and in-situ soil moisture measurements, especially for small-scale modelling studies.

CRediT authorship contribution statement

Shaochun Huang: Conceptualization, Methodology, Software, Writing – original draft. **Stephanie Eisner:** Data curation, Writing – review & editing. **Ingjerd Haddeland:** Data curation, Writing – review & editing. **Zelalem Tadege Mengistu:** Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Shaochun Huang, Stephanie Eisner, Ingjerd Haddeland reports financial support was provided by The Research Council of Norway.

Acknowledgement

The authors would like to thank The Research Council of Norway for funding the IMPRINT (The human imprint on land-atmosphere exchange in high latitudes) project no. 295128.

Data availability

HySN2018v2005ERA5 DOI: 10.5281/zenodo.5947547.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2022.127895.

References

Abbaspour, K.C., Vaghefi, S.A.S., Yang, H., Srinivasan, R., 2019. Global soil, landuse, evapotranspiration, historical and future weather databases for SWAT Applications. Sci. Data 6. https://doi.org/10.1038/s41597-019-0282-4.

Ahlstrøm, A., Bjørkelo, K., Frydenlund, J., 2014. AR5 klassifikasjonssystem - klassifikasjon av arealressurser. Skog og landskap 38.

Andreadis, K.M., Storck, P., Lettenmaier, D.P., 2009. Modeling snow accumulation and ablation processes in forested environments. Water Resour. Res. 45 https://doi.org/10.1029/2008wr007042.

Arnold, J.G., Allen, P.M., Bernhardt, G., 1993. A Comprehensive Surface-Groundwater Flow Model. J. Hydrol. 142 (1–4), 47–69.

Arnold, J.G., Williams, J.R., Nicks, A.D., Sammons, N.B., 1990. SWRRB - A Basin Scale Simulation Model for Soil and Water Resources Management. Texas A&M University Press, College Station, Texas.

Baritz, R., Seufert, G., Montanarella, L., Van Ranst, E., 2010. Carbon concentrations and stocks in forest soils of Europe. Forest Ecol Manag 260 (3), 262–277. https://doi.org/10.1016/j.foreco.2010.03.025.

Batjes, N.H., 2003. A taxotransfer rule-based approach for filling gaps in measured data in primary SOTER databases. ISRC-World Soil Information, Wageningen.

Batjes, N.H., 2016. Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. Geoderma 269, 61–68. https://doi.org/10.1016/j.geoderma.2016.01.034.

Batjes, N.H., Ribeiro, E., van Oostrum, A., 2020. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). Earth Syst. Sci. Data 12 (1), 299–320. https://doi.org/10.5194/essd-12-299-2020.

Beck, H.E., Pan, M., Miralles, D.G., Reichle, R.H., Dorigo, W.A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R.M., van Dijk, A.I.J.M., Du, J.Y., Kimball, J. S., Vergopolan, N., Wood, E.F., 2021. Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors. Hydrol. Earth Syst. Sci. 25 (1), 17–40. https://doi.org/10.5194/hess-25-17-2021.

Blyverket, J., Hamer, P.D., Schneider, P., Albergel, C., Lahoz, W.A., 2019. Monitoring Soil Moisture Drought over Northern High Latitudes from Space. Remote Sens. Basel 11 (10). https://doi.org/10.3390/rs11101200.

Boone, A., Habets, F., Noilhan, J., Clark, D., Dirmeyer, P., Fox, S., Gusev, Y., Haddeland, I., Koster, R., Lohmann, D., Mahanama, S., Mitchell, K., Nasanova, O., Niu, G.-Y., Pitman, A., Polcher, J., Shmakin, A.B., Tanaka, K., van den Hurk, B., Verant, S., Verseghy, D., Viterbo, P., Yang, Z.-L., 2004. The Rhone-Aggregation Land Surface Scheme Intercomparison Project: An Overview. J. Clim. 17 (1), 187–208.

Bright, R.M., Eisner, S., Lund, M.T., Majasalmi, T., Myhre, G., Astrup, R., 2018. Inferring Surface Albedo Prediction Error Linked to Forest Structure at High Latitudes. J. Geophys. Res. [Atmos.] 123 (10), 4910–4925.

Choi, J., Won, J., Lee, O., Kim, S., 2021. Usefulness of Global Root Zone Soil Moisture Product for Streamflow Prediction of Ungauged Basins. Remote Sens-Basel 13 (4). https://doi.org/10.3390/rs13040756. ARTN 756.

Dai, Y.J., Wei, S.G., Nan, W., Xin, Q.C., Hua, Y., Zhang, S.P., Liu, S.F., Lu, X.J., Wang, D. G., Yan, F.P., 2019. A review of the global soil property maps for Earth system models. Soil-Germany 5 (2), 137–158. https://doi.org/10.5194/soil-5-137-2019.

de Lannoy, G.J.M., Koster, R.D., Reichle, R.H., Mahanama, S.P.P., Liu, Q., 2014. An updated treatment of soil texture and associated hydraulic properties in a global land modeling system. J. Adv. Model Earth Syst. 6 (4), 957–979. https://doi.org/10.1002/2014ms000330.

Erlandsen, H.B., Beldring, S., Eisner, S., Hisdal, H., Huang, S.C., Tallaksen, L.M., 2021. Constraining the HBV model for robust water balance assessments in a cold climate. Hydrol. Res. 52 (2), 356–372. https://doi.org/10.2166/nh.2021.132.

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 Syst. Sci. Data 11 (2), 797–821. https://doi.org/10.5194/essd-11-797-2019.

FAO, 2003. The Digitized Soil Map of the World Including Derived Soil Properties (version 3.6), FAO, Rome, Italy.

FAO/IIASA/ISRIC/ISSCAS/JRC, 2012. Harmonized World Soil Database (version 1.2). In: FAO (Ed.), Rome, Italy and IIASA, Laxenburg, Austria.

Franchini, M., Pacciani, M., 1991. Comparative analysis of several conceptual rainfall-runoff models. J. Hydrol. 122 (1-4), 161–219.

Gjertsen, A.K., Nilsen, J.E., 2012. SAT-SKOG: Et skogkart basert på tolking av satellittbilder, Skog og landskap, rapport nr. 23/2012, 54 pp. http://hdl.handle.net/11250/2453917.

Global Soil Data Task Group: Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS), Data set, International Geosphere-Biosphere Programme – Data and Information System, available at: http://www.daac.ornl.gov from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, 2000.

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., van der Schalie, R., Wigneron, J. P., Wagner, W., 2020. Validation practices for satellite soil moisture retrievals: What are (the) errors? Remote Sens. Environ. 244, 111806 https://doi.org/10.1016/j.rse.2020.111806.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., Dorigo, W., 2019. Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. Earth Syst. Sci. Data 11 (2), 717–739. https://doi.org/10.5194/essd-11-717-2019.

Guillod, B.P., Davin, E.L., Kundig, C., Smiatek, G., Seneviratne, S.I., 2013. Impact of soil map specifications for European climate simulations. Clim Dynam 40 (1–2), 123–141. https://doi.org/10.1007/s00382-012-1395-z.

Haddeland, I., Matheussen, B.V., Lettenmaier, D.P., 2002. Influence of spatial resolution in a macroscale hydrologic model. Water Resour. Res. 38(7), Art. No, 1124. https://doi.org/10.1029/2001WR000854.

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.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., Bond-Lamberty, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS One 12 (2), e0169748.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q J Roy Meteor Soc 146 (730), 1999–2049.

Huang, S.C., Krysanova, V., Hattermann, F.F., 2013. Projection of low flow conditions in Germany under climate change by combining three RCMs and a regional hydrological model. Acta Geophys. 61 (1), 151–193. https://doi.org/10.2478/s11600-012-0065-1.

Huang, S.C., Kumar, R., Florke, M., Yang, T., Hundecha, Y., Kraft, P., Gao, C., Gelfan, A., Liersch, S., Lobanova, A., Strauch, M., van Ogtrop, F., Reinhardt, J., Haberlandt, U.,

Krysanova, V., 2017. Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide. Clim. Change 141 (3), 381–397. https://doi.org/10.1007/s10584-016-1841-8.
Kearney, M.R., Maino, J.L., 2018. Can next-generation soil data products improve soil moisture modelling at the continental scale? An assessment using a new microclimate package for the R programming environment. J. Hydrol. 561, 662–673. https://doi.org/10.1016/j.jhydrol.2018.04.040.
Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F., Kundzewicz, Z.W., 2018. How the performance of hydrological models relates to credibility of projections under climate change. Hydrol. Sci. J. 63 (5), 696–720. https://doi.org/10.1080/02626667.2018.1446214.
Krysanova, V., Meiner, A., Roosaare, J., Vasilyev, A., 1989. Simulation modelling of the coastal waters pollution from agricultural watersheds. Ecol. Model. 49 (1-2), 7–29.
Krysanova, V., Müller-Wohlfeil, D.-I., Becker, A., 1998. Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. Ecol. Modell. 106 (2-3), 261–289.
Kværnø, S.H., Haugen, L.E., 2011. Performance of pedotransfer functions in predicting soil water characteristics of soils in Norway. Acta Agriculturae Scandinavica, Section B - Soil Plant Sci. 61 (3), 264–280.
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.
Livneh, B., Kumar, R., Samaniego, L., 2015. Influence of soil textural properties on hydrologic fluxes in the Mississippi river basin. Hydrol. Process. 29 (21), 4638–4655. https://doi.org/10.1002/hyp.10601.
Lussana, C., Tveito, O.E., Dobler, A., Tunheim, K., 2019. seNorge_2018, daily precipitation, and temperature datasets over Norway. Earth Syst. Sci. Data 11 (4), 1531–1551. https://doi.org/10.5194/essd-11-1531-2019.
Ma, H.L., Zeng, J.Y., Chen, N.C., Zhang, X., Cosh, M.H., Wang, W., 2019. Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations. Remote Sens Environ 231. https://doi.org/10.1016/j.rse.2019.111215. ARTN 111215.
Magnusson, J., Eisner, S., Huang, S., Lussana, C., Mazzotti, G., Essery, R., Saloranta, T., Beldring, S., 2019. Influence of Spatial Resolution on Snow Cover Dynamics for a Coastal and Mountainous Region at High Latitudes (Norway). Water Resour. Res. 55 (7), 5612–5630. https://doi.org/10.1029/2019wr024925.
Maidment, D.R., 1993. Handbook of hyrology. McGraw-Hill Inc., New York.
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.
Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernandez-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10 (5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017.
McBratney, A.B., Santos, M.L.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.
Mukundan, R., Radcliffe, D.E., Risse, L.M., 2010. Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. J. Soil Water Conserv. 65 (2), 92–104. https://doi.org/10.2489/jswc.65.2.92.
Orth, R., Seneviratne, S.I., 2013. Propagation of soil moisture memory to streamflow and evapotranspiration in Europe. Hydrol Earth Syst. Sci. 17 (10), 3895–3911. https://doi.org/10.5194/hess-17-3895-2013.
Pellarin, T., Calvet, J.C., Wagner, W., 2006. Evaluation of ERS scatterometer soil moisture products over a half-degree region in southwestern France. Geophys. Res. Lett. 33 (17). Artn L17401 10.1029/2006gl027231.
Pelletier, J.D., Broxton, P.D., Hazenberg, P., Zeng, X.B., Troch, P.A., Niu, G.Y., Williams, Z., Brunke, M.A., Gochis, D., 2016. A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. J. Adv. Model Earth Syst. 8 (1), 41–65. https://doi.org/10.1002/2015ms000526.
Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Reistad, M., Breivik, O., Haakenstad, H., Aarnes, O.J., Furevik, B.R., Bidlot, J.R., 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.
Renzullo, L.J., van Dijk, A.I.J.M., Perraud, J.M., Collins, D., Henderson, B., Jin, H., Smith, A.B., McJannet, D.L., 2014. Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. J. Hydrol. 519, 2747–2762. https://doi.org/10.1016/j.jhydrol.2014.08.008.
Riley, H., 1996. Estimation of physical properties of cultivated soils in southeast Norway from readily available soil information. Norwegian J. Agric. Sci. (Supplement No. 25), 1–51.
Ruan, G., Langsholt, E., 2017. Rekalibrering av flomvarslingas HBV-modeller med inndata fra seNorge, versjon 2.0. 71, Norges vassdrags- og energidirektorat, Oslo, Norway.
Sangrey, D., Harrop-Williams, K., Kaliber, J., 1984. Predicting ground-water response to precipitation. ASCE J. Geotech. Eng. 110 (7), 957–975.
Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Sci. Rev. 99 (3–4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004.
Shangguan, W., Dai, Y.J., Duan, Q.Y., Liu, B.Y., Yuan, H., 2014. A global soil data set for earth system modeling. J. Adv. Model Earth Syst. 6 (1), 249–263. https://doi.org/10.1002/2013ms000293.
Sheshukov, A.Y., Daggupati, P., Mankin, K.R., 2011. High spatial resolution soil data for watershed modeling: 2. assessing impacts on watershed hydrologic response. Nat. Environ. Sci. 2 (2), 32–41.
Sveistrup, T.E., 1984. Retningslinjer for beskrivelse av jordprofil. (Guideline for soil profile description.). Jord og Myr 8 (2), 47.
Tian, S.Y., Renzullo, L.J., van Dijk, A.I.J.M., Tregoning, P., Walker, J.P., 2019. Global joint assimilation of GRACE and SMOS for improved estimation of root-zone soil moisture and vegetation response. Hydrol. Earth Syst. Sci. 23 (2), 1067–1081. https://doi.org/10.5194/hess-23-1067-2019.
Tifafi, M., Guenet, B., Hatte, C., 2018. Large Differences in Global and Regional Total Soil Carbon Stock Estimates Based on SoilGrids, HWSD, and NCSCD: Intercomparison and Evaluation Based on Field Data From USA, England, Wales, and France. Global Biogeochem. Cy 32 (1), 42–56. https://doi.org/10.1002/2017gb005678.
Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour. Res. 43 (1). Artn W01413 10.1029/2005wr004723.
Wagner, W., Lemoine, G., Rott, H., 1999. A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sens. Environ. 70 (2), 191–207.
Weynants, M. et al., 2013. European HYdropedological Data Inventory (EU-HYDI). JRC81129, Luxembourg.
Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90 (3-4), 169–185.
Ye, X.C., Zhang, Q., Viney, N.R., 2011. The effect of soil data resolution on hydrological processes modelling in a large humid watershed. Hydrol. Process. 25 (1), 130–140. https://doi.org/10.1002/hyp.7823.