The impact of initial conditions on convection-permitting simulations of a flood event over complex mountainous terrain

Lu Li[1], Marie Pontoppidan[1], Stefan Sobolowski[1], Alfonso Senatore[2]

[1]NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Jahnebakken 5, 5008 Bergen, Norway
5 [2]Department of Environmental and Chemical Engineering, University of Calabria, Italy

Correspondence to: Lu Li (luli@norceresearch.no)

Abstract. Western Norway suffered major flooding after 4 days of intense rainfall during the last week of October 2014. While events like this are expected to become more frequent and severe under a warming climate. Convection-permitting scale models are showing their skills in capturing their dynamics. Nevertheless, several sources of uncertainty need to be taken into 10 account, among which the impact of initial conditions on precipitation pattern and discharge, especially over complex, mountainous terrain. In this paper, the Weather Research and Forecasting model Hydrological modelling system (WRF-Hydro) is applied at a convection-permitting scale and its performance is assessed in western Norway for the aforementioned flood event. The model is calibrated and evaluated using observations and benchmarks obtained from the Hydrologiska Byråns Vattenbalansavdelning (HBV) model. The calibrated WRF-Hydro model performs better than the simpler conceptual HBV 15 model, especially in areas with complex terrain and poor observational coverage. The sensitivity of the precipitation pattern and discharge to poorly constrained elements such as spinup time and snow conditions is then examined. The results show that: (1) overall, the convection-permitting WRF-Hydro simulation captures the precipitation pattern/amount, the peak flow volume and the timing of the flood event; (2) precipitation is not overly sensitive to spinup time, while discharge is slightly more sensitive due to the influence of soil moisture, especially during the pre-peak phase; (3) the idealized snow depth 20 experiments show that a maximum of 0.5 m of snow is converted to runoff irrespective of the initial snow depth and that this snowmelt contributes to discharge mostly during the rainy and the peak flow periods. Though further targeted experiments are needed, this study suggests that snow-cover intensifies the extreme discharge instead of acting as a sponge, which implies that future rain-on-snow events may contribute to higher flood risk.

1 Introduction

25 Heavy rainfall, with local amounts exceeding 350 mm, fell over the coastal and mountainous areas of western Norway between 26 - 29 October 2014. The event caused widespread flooding, with 16 stations registering discharge above the 50-year flood threshold (Langsholt et al., 2015). The severe flooding of many regional river systems destroyed infrastructure, houses and isolated towns. Overall the damage exceeded 131.6 million Norwegian kroner (~ 16.4 million USD) (Dannevig et al., 2016).

A list of reports (in Norwegian), e.g. "October flood in western Norway 2014" (Dannevig et al., 2016) and "The flood in western Norway October 2014" (Lansholt et al., 2015) from Norwegian Water Resources and Energy Directorate (NVE) were produced, which documented the rainfall and discharge records and societal impacts. For flood hazards such as this, it is a challenge to forecast/hindcast the hydrological response due to the complex terrain and the events' complex spatial and temporal characteristics. However, with extreme precipitation over this region projected to increase significantly over the coming decades (e.g. Hanssen-Bauer et al., 2017) the need to reliably reproduce such events is high for both climate researchers and operational professionals.

1.1 Norwegian flood types, changes & rain-on-snow

The complex and varying terrain of Norway divides the country into different climatic zones and flood-generating processes vary by regions. For example, northern and eastern Norway are mainly prone to spring snowmelt floods while southern and western Norway are dominated by rain induced floods (Vormoor et al., 2015). According to recent studies, snowmelt generated floods have decreased and shifted earlier in the spring in recent decades (Vormoor et al., 2016; Pall et al. 2018). At the same time, rain dominated floods are increasing in frequency (Vormoor et al., 2016). This is consistent with observed increases in precipitation (Dyrrdal et al., 2012) and streamflow (Stahl et al., 2010; Wilson et al., 2010), and the trends are projected to continue into the future (Hanssen-Bauer et al., 2017; Sorteberg et al., 2018). Though temperatures are increasing, much of the winter precipitation in inland catchments will continue to fall as snow until at least the mid-century (Hanssen-Bauer et al., 2017). However, as temperature rises many of these catchments may experience more rain-on-snow events. Warm, and often windy conditions during such events can cause substantial additional snowmelt, which can exacerbate an already dangerous flooding event (Marks et al., 1998, 2001). In fact, for many catchments in the world, such as the western US, rain-on-snow events are important for prediction of flood responses and risk (Berghuijs et al., 2016; Musselman et al., 2018). While earlier snowmelt is decreasing the frequency of such events in the late spring and at low altitudes, both the magnitude and frequency of rain-on-snow events are increasing during winter in central Europe (Freudiger et al., 2014), and likely also in Norway. Pall et al. (2018) used a high-resolution (1 km) seNorge data set to construct climatology of rain-on-snow occurrence in the mainland Norway for recent decades. They found an increase in rain-on-snow events in high-elevation areas across the mainland in winter-spring. Given the dependence of floods in Norway on a complex interplay between variations in elevation, temperature gradients (e.g. between land and ocean), orographic interactions, existing snow and soil moisture distributions, etc. it is critical to run models (either dynamical or statistical) at resolutions that can capture this complexity. But this requirement presents challenges of its own. Despite this, the rain versus snowmelt contribution to the flow can be important in determining the flood generation processes for Norway and can be particularly sensitive to the vertical temperature gradient.

1.2 Forcing data and convection-permitting modelling

In order to improve our scientific understanding as well as predictions and projections of flooding, high quality meteorological forcing data is crucial. A lack of detailed precipitation records that accurately represent spatial and temporal variability at both the basin and regional scales presents well-known challenges to hydrological modelling. In mountainous areas, like western Norway, where precipitation is strongly influenced by the terrain, spatial patterns of precipitation are not well captured by either sparse gauge data or gridded precipitation datasets (e.g. satellite-based products or high-resolution interpolation-based datasets). High-resolution, convection-permitting modelling has exhibited great promise in addressing these issues and has potential as a powerful tool for hydrological prediction (Prein et al., 2015, 2016, 2017; Smiatek et al., 2016; Kendon et al., 2017; El-Samra et al., 2018; Poschlod et al., 2018; Avolio et al., 2019). Pontoppidan et al. (2017) investigated the importance of kilometre-scale resolution on the aforementioned flooding event in October 2014 over Western Norway and found that convection-permitting simulations (~3 km grid spacing) from the Weather Research and Forecasting (WRF) model substantially improved the representation of precipitation compared to a coarser 9 km grid spacing simulation. This improvement was seen both in terms of absolute values and spatial-temporal distribution. The largest improvement simulation with the resolution jump from parameterized to explicitly resolved convection (e.g., 9 km to 3 km over western Norway). Several previous studies over other regions also demonstrate the added-value of convection permitting modelling for extreme weather impact studies in regions with complex terrain. For example, Maussion et al. (2011) showed improved representation of precipitation in a convection-permitting (2 km) simulation when compared to satellite products over the Himalayan region. El-Samra et al. (2018) suggest that downscaling over complex terrain requires a horizontal grid resolution of 3 km or higher in order to improve the forecasting of mean and extreme temperatures and capture the orographic precipitation climatology. Conversely, coarse resolution (~ 9 km) simulations miss the impact of orography on temperature and precipitation. Additionally, the studies of Rasmussen et al. (2011, 2014) found that a spatial and temporal depiction of snowfall that is adequate for water resource management over the Colorado Headwaters regions can only be achieved with the appropriate choice of model grid spacing and parameterizations. The modeling systems that are capable of accurately depicting the atmosphere at these scales now increasingly incorporate other regional system components such as crops, urban features and, of most relevance to the present study, hydrology.

1.3 A dynamical hydrometeorological model: WRF-Hydro modelling system

The Weather Research and Forecasting Model Hydrological modeling system (WRF-Hydro) is a model coupling framework designed to link multi-scale process models of the atmosphere and terrestrial hydrology (Gochis et al., 2018). It runs both in fully-coupled (two-way) or uncoupled (one-way, from atmosphere to land) modes and is intended to serve as both a hydrometeorological prediction system and a research tool. The system has been applied in studies around the world (e.g. Senatore et al. 2015; Givati et al. 2016; Arnault et al., 2016; Xiang et al., 2017; Naabil et al., 2017; Verri et al., 2017; Lin et

al., 2018; Rummler et al., 2019). It is currently in use operationally as a key component of the United States national water model where it expands the number streamflow forecast points from ~3600 points to ~2.7 million river reaches (https://water.noaa.gov/about/nwm). WRF-Hydro has also been applied in Africa (Arnault et al., 2016; Kerandi et al., 2018), in the Himalayas (Li et al., 2017), in Italy (Verri et al., 2017; Senatore et al., 2015) and in Eastern Alps (Rummler et al., 2019) with promising results and shows potential for use in runoff forecasting, water resource planning and climate changes impact assessments. However, despite application across a diverse array of catchments and research questions, the system has yet to be evaluated for a case in Norway.

There are still challenges to discharge prediction by WRF-Hydro and the performance varies across geographic regions and climate. For example, it simulated flood events in the Black Sea region fairly well if both model calibration and WRF data assimilation were performed jointly, while the streamflow obtained with raw WRF precipitation was in general very poor (Yucel et al., 2015). It also simulated a full annual cycle of the Crati River basin in southern Italy with Nash-Sutcliffe Efficiency (NSE) of 0.8 using observed precipitation while only achieved an NSE of 0.27 using simulated precipitation (a perfect model result of NSE is 1.0 and NSE value of 0 indicates that the model predictions are as accurate as the mean of the observed data) (Senatore et al., 2015). Naabil et al. (2017) applied WRF-Hydro in a test case over west Africa for water resources management, and found that further improvements via proper model calibration and consideration of the effects of model biases in the dam level were recommended, although model captured the attributes of the streamflow. Furthermore, Verri et al. (2017) demonstrated that the performance of WRF-Hydro was severely affected by various components including simulated precipitation, initial conditions and also the calibration / validation of discharge hydrography. In Texas, WRF-Hydro has shown promise as a forecast tool, but suffers from poor prediction skill in areas with human altered flows, in which both the surface runoff and the base flow are underpredicted (Lin et al., 2018). Additional studies also noted the sensitivity of WRF-Hydro to the initial conditions (spinup time) (Roman-Cascon et al., 2016; Bonekamp et al., 2018; Verri et al., 2017). In order to obtain a stable WRF-Hydro simulation, a spinup period is required, which depends on the quality of the model input and soil data. Therefore, the impact of the spinup time needs to be assessed on per-case basis as it likely depends on local conditions.

1.4 Objectives of the paper

Due to the traditional separation of hydrological and atmospheric modelling communities, significant gaps exist in our knowledge of the full-chain responses to hydrometeorological extremes, from circulation/transport of moisture to precipitation to discharge. WRF-Hydro is designed to link across these components and their characteristic scales to provide a modelling framework that can address these gaps (Gochis et al., 2018). It enables improved simulation of land surface hydrology and energy states and fluxes at high spatial resolution (typically 1 km or less). It can be used in either "offline" (uncoupled to the atmospheric component of the model) or "fully-coupled" modes (the hydrological model components have two-way interactions with the atmospheric component) (Gochis et al., 2015).

In this study, we employ WRF-Hydro in western Norway to investigate meteorological and hydrological processes driving the October 2014 flooding. To our knowledge, this is the first study using a complete meteorological-hydrological modelling approach to characterize a precipitation induced extreme flooding event in Norway. The causal mechanisms and evolution of this particular flood event are examined. In addition, we explore the sensitivity of the discharge to different initial conditions such as soil and snow.

The work is built upon the study of Pontoppidan et al. (2017) for the simulation of the meteorological processes and the hydrological impact. As such, an 'offline' ('uncoupled') configuration for the WRF-Hydro model is chosen. This is because we primarily aim to understand the flood event in the context of its hydrological response to the weather forcing and land surface conditions. Feedbacks between the atmosphere and land, though important, generate second order effects that likely have only small impact within such a short duration event. Also, the offline mode of WRF-Hydro system is preferable for our study as it provides a clearer interpretation of the results, identification of uncertainties in the water budget and assessment of sensitivity to critical parameters in the atmospheric and hydrological components (Li et al., 2017).

The remainder of the paper is structured as follows: after the introduction, a description of the study area and data is presented, followed by methods, including a description of the WRF-Hydro setup and experiment design, model calibration and benchmark evaluation. Results concentrate on the model calibration and benchmark evaluation, precipitation evaluation, and the impacts of initialization (spinup time) and prescribed snow cover are examined. Finally, the main conclusions are presented.

2 Study area and data

2.1 Study area

Our four study catchments are located in western Norway, where the landscape is dominated by steep orography and complex terrain due to the fjords and elevation varies from sea level to more than 2400 m (Figure 1). The complex terrain both enhances the precipitation and generates large local differences in the precipitation distribution (e.g. Reuder et al. 2007, Pontoppidan et al 2017). Norway is positioned at the exit region of the North Atlantic storm track, which brings low-pressure systems and associated frontal precipitation towards the west coast on a regular basis during autumn and winter. Western Norway is the wettest part of the country (Hanssen-Bauer and Førland, 2000) and annual precipitation exceeds 3000 mm in several places; but there is also high spatial variability. For example, Kvamskogen-Jonshøgdi (60.389 N and 5.964 E) records 3151 mm while Vossevangen (60.625 N and 6.426 E), which is only 36.6 km away, only receives 1280 mm (MET Norway, 2015) (see in Figure 1).

2.2 Hydro-meteorological conditions

October 2014 was wetter than usual in western Norway. The situation was maintained by an atmospheric river and the passage of multiple frontal systems with moderate to heavy precipitation. Two days before the flooding event, on the 26th of October, a low-pressure system with associated fronts passed over western Norway and delivered considerable amounts of precipitation. A cold front passed the area at midnight on the 27th, advecting colder and drier air into the area for a short period. Simultaneously, a disturbance over Scotland developed and moved towards Norway, leaving western Norway in the warm sector of an intensifying low-pressure system. Once again large amounts of precipitation fell from midday on the 27th to early evening on the 28th. The associated cold front passed the Bergen area in the afternoon and the precipitation intensity decreased with its passage. Due to the already saturated soil and several days with more or less continuous rainfall, the flood peaked in the Voss area in early evening of the 28th (Pontoppidan et al., 2017).

According to the NVE report made by Langsholt et al. (2015), there was shallow snow cover in high-altitude areas east of the catchments. In Voss however, where our four catchments are located, there was no snow in the snow depth water equivalent maps released by NVE. These maps are made from model simulations based on NVE snow observations. Discharge in each of the study catchments was over the 50-year return level. On the 29th of October, the daily discharge record held since 1892 was broken at the Bulken station, located at the outlet of Vangsvatnet (Langsholt et al., 2015).

2.3 Observational data

We use 43 precipitation gauges from the Norwegian Meteorological Institute (MET Norway) situated in and around the catchments, with either hourly or daily precipitation data. Typically rain gauges in Norway are deployed at low elevations and in valleys resulting in skewed precipitation distributions. To rectify this, 11 HOBO rain gauges, which provide hourly data, were deployed at higher elevations in a transect from the coast to inland (Pontoppidan et al., 2017). A table with station details can be found in Pontoppidan et al. (2017). Catchment averaged precipitation is calculated as a mean of the rain gauges positioned within the affected catchment. Four discharge stations from NVE are used for WRF-Hydro model discharge calibration and validation (see Table 1). It should be noted that the drainage basin of the Bulken catchment includes the catchments of Kinne and Myrkdalsvatn. In addition, four precipitation gauges from MET Norway, which are the nearest station in the four basins, are chosen for precipitation evaluation in section 4.3 (Table 1). Figure 1 shows the locations of the four catchments and the measurement sites of rainfall and discharge gauges.

3 Methods

3.1 WRF domain design

175 The Advanced Research WRF (WRF-ARW) model version 3.9.1 is set up with two nested domains with spatial resolution of 9 km and 3 km (Figure 1). The lateral boundaries are forced with the 6-hourly ERA-interim reanalysis with a spatial resolution of 0.75 degrees (Dee et al. 2011). The sea surface temperatures (SST) are also updated every 6 hours. The model is run with 40 vertical levels in all domains.

The choice of the microphysical scheme is important for precipitation. Previous studies of mountain precipitation using WRF

180 have shown that the Thompson microphysical scheme (Thompson et al. 2008) performs well (Collier et al. 2013; Maussion et al. 2014; Rasmussen et al. 2011, 2014, Li et al. 2017), especially in areas with mixed hydrometeors because it computes cloud water, rain water, snow, graupel and ice. The scheme was also successfully used in a previous study on this specific event (Pontoppidan et al. 2018). The grid spacing in the outer domain is in the so-called "gray zone" (5-10 km) where convection may or may not be explicitly resolved; therefore, we tested the impact of the convection parameterization on precipitation. The

185 results showed negligible differences between simulations with the convection scheme on and off. Here we present the results from the simulations with the convection parameterization turned off. The Yonsei University scheme (Hong et al., 2006) is used for the planetary boundary layer, the RRTM scheme for long wave radiation (Mlawer et al., 1997) and the RRTMG scheme for shortwave radiation (Iacono et al., 2008). The Noah Land Surface model ('Noah LSM', Mitchell et al. 2001) is used for surface scheme, which has a bulk layer simple canopy and snow model. In the Noah LSM, the snow cover area

190 fraction within a model grid is determined as a function of snow water equivalent (SWE) using a generalized snow depletion curve. When snow is on the ground, the model considers a bulk snow-soil-canopy layer and computes surface temperature at each time step. The snow surface temperature for the snowpack is estimated in two steps. Firstly, the energy balance between the snowpack, top-soil layer and the overlying air is calculated to obtain an intermediate temperature. This temperature can rise above freezing even when the model grid is fully covered with snow. Secondly, the effective temperature is adjusted by

195 accounting for the fractional snow cover on the ground (Livneh et al., 2010).

Additional configuration details include a model time-step of 18 seconds over the inner domain and the use of spectral nudging to keep the large-scale flow consistent with the driving ERA-Interim reanalysis. This approach proved to be useful when reproducing extreme precipitation events due to the better resolved synoptic scale features over North Atlantic (Heikkilä et al., 2011). Spectral nudging (Radu et al., 2008) is only applied in the outer domain leaving the model free to create its own

200 structures in the inner domain. In the present case, nudging is only applied above the boundary layer and only on wavelengths longer than 585 km.

3.2 WRF-Hydro modelling system

Version 3.0 of the WRF-Hydro modelling system is used in the study. A comprehensive description of the model system can be found in Gochis et al. (2015). In our study, the saturated subsurface overflow routing, surface overland flow routing, channel routing and base-flow modules are activated. The overland flow routing adopts a 2-D diffusive wave formulation (Julien et al., 1995) and the channel routing is calculated by a 1-D variable time-stepping diffusive wave formulation. In addition, a bucket model for base-flow is used where a groundwater reservoir with conceptual depth and a related conceptual volume is associated. A few lakes in Bulken catchment are not considered in this study due to lack of data.

WRF-Hydro is set up to run offline using the WRF atmospheric simulations as input (see Introduction). The subgrid routing processes are executed at a 300 m grid spacing and the surface physiographic files are prepared by ARCGIS 10.6 (Sampson et al. 2018). The physiographic file includes high-resolution terrain grids specifying the topography, channel grid, flow direction, stream order (for channel routing), ground-water basin mask and the position of stream gauging stations which are the outlets for water routing out across the landscape (Gochis et al., 2015). There are four stream orders in the network of the study catchments shown in Figure 2.

3.3 Model calibration

Two-step calibration of WRF-Hydro is performed in the study. First, we select the most sensitive parameters from a wide range of parameters. These include: the saturation soil conductivity (in SOILPARM.TBL), optimum transpiration air temperature (in VEGPARM.TBL) and infiltration parameter (in the surface runoff parameterization of GENPARM.TBL), Manning roughness coefficients (in the channel routing of CHANPARM.TBL), the groundwater bucket model exponent (in the groundwater bucket model of GWBUCKPARM.TBL), surface flow roughness scaling factor (OVROUGHRTFAC) and the surface retention depth (RETDEPRT) (Yucel et al., 2015; Li et al., 2017). Second, three parameters, which are particularly sensitive, are tuned using the auto-calibration Parameter Estimation Tool (PEST http://www.pesthomepage.org): two infiltration parameters, i.e. REFDK_DATA (refdk) and REFKDT_DATA (refkdt), which are important for surface runoff, and the Manning routing coefficients (mn01). The offline model is then forced by meteorological output data and calibrated based on the observed discharge in the Svartavatn catchment. The remaining three catchments (i.e. Bulken, Kinne and Myrkdalsvatn) are used for validation and evaluation of the parameters' transferability. The simulations are initialized on 1 September 2014 and run until 1 November 2014. In order to remove the impact of initialization we use the first 30 days as spinup in the model calibration. The best parameter set is then chosen based on the Nash-Sutcliffe efficiency (NSE) coefficient (Nash and Sutcliffe, 1970). Two more indices, bias and root mean square error (RMSE) are also used for validation. A perfect model would have an NSE value of 1 and bias and RMSE values equal to 0. In addition, the correlation coefficient matrix of calibrated parameters is also estimated by the PEST method (Doherty, 2015). It tells which two parameters might be linearly dependent (if the correlation coefficient is greater than 0.8).

3.4 Benchmark evaluation approach

A simple bucket-type Hydrologiska Byråns Vattenbalansavdelning (HBV) light model was used as a benchmark for model comparison and evaluation (Seibert and Vis, 2012; Seibert et al., 2018). The HBV light is an offshoot of the HBV model developed in the 1970s by the Swedish Meteorological and Hydrological Institute (SMHI). It consists of four main routines, i.e., snow-, soil-, routing- and response routine and simulates daily discharge using daily precipitation, temperature and potential evapotranspiration (Seibert and Vis, 2012). Its strength lies in the relatively low requirements for input data and the limited number of parameters (Rusli et al., 2015). Here, the calibrated HBV streamflow is used as upper benchmark (R_{upper}) and two alternatives are then used as lower benchmark (R_{lower}), one generated from the mean streamflow from 1000 random parameter sets ($R_{lower/random}$) and another from the regionalization parameter set from other nearby catchments ($R_{lower/regional}$). The catchment averaged daily precipitation, temperature and potential evaporation from WRF are used as input in the HBV model simulation. To maintain consistency with the WRF-Hydro modeling, the HBV simulations are also initiated on 1 September 2014 and run until 1 November 2014 with the first 30 days used as spinup. The performance measure for the benchmark evaluation is the NSE.

3.5 Initialisation experiments

Previous studies found that spinup time influences the initial conditions such as the soil moisture content and therefore the latent heat flux, which in turn influences the precipitation (Kleczek et al., 2014; Bonekamp et al., 2018; Verri et al., 2017). Jankov et al. (2007) suggested the spinup time should be at least 12 h to prevent instabilities in WRF, but the recommended length most likely depends on the input quality and soil fields (Kleczek et al., 2017). For example, Bonekamp et al., (2018) found that precipitation is extremely sensitive to the spinup time in summer, with the best performance coming with 24 hours of spinup, while does not show a clear trend with increasing spinup time over 24 hours. For our study, it is not known *a priori* how the model simulation will be effected by the spinup time. So we conduct experiments with different spinup times ranging from 1 day to 26 days, and investigate the influence of spinup time on the amount of precipitation, soil moisture and outlet discharge of the extreme event in the study. An overview of the initialization experiments performed in the paper is given in Table 2. The evaluation period is 23 - 31 October 2014, and includes a minor peak flow on the 24th of October before the major peak flows on 26th and 28th.

3.6 Prescribed snow cover experiments

In the October 2014 flood event, temperatures in the mountains were above freezing and the ground was bare. In other words, there was no layer of snow to act as a sponge and potentially affect the discharge. In a future warmer climate, however, rain-on-snow events are likely to increase, especially in mountainous areas of Norway (Vormoor et al., 2016). However, the potential impact of snow conditions on extreme flows is not well known. Therefore, we construct a series of hypothetical

experiments for a primary check on this impact. The results can be helpful for filling this knowledge gap and dictating the flood generation processes for Norway, although we know this hypothetical case most likely will increase in eastern and northern Norway instead of western Norway.

In the study, we perform two types of snow experiments: (1) different uniform snow depths are applied over the entire study area (i.e., 0.1 m, 0.5 m, 1 m and 2 m) and (2) 1 m of snow is imposed above certain elevations (i.e. 400 m, 600 m and 800 m above sea level (a.s.l.)). The experiments are all performed with the calibrated parameter set. More details can be found in Table 3. We are mainly interested in evaluating the precipitation-snowmelt timing and snowmelt augmentation of the peak flow, if any. Therefore, we apply the prescribed snow cover fields in the restart file on the 25th of October 2014, which is from the 26-day spinup experiment. The area-elevation distribution in the four selected catchments is shown in Table 4. The Kinne and Myrkdalsvatn catchments are dominated by higher elevations with 48 % and 44 % of the area above 1000 m a.s.l., respectively, compared to 36 % and 9 % from Bulken and Svartavatn (Table 4).

4 Results

4.1 WRF-Hydro discharge calibration

Since calibration is computationally demanding we calibrate WRF-Hydro based on the discharge of Svartavatn, which is the smallest catchment in the study region. The remaining three catchments are used for model validation. The calibration and validation results are shown in Table 5. The Nash-Sutcliffe-Efficiency coefficient (NSE) of daily discharge increases from 0.41 to 0.86, while the Bias and RMSE decrease from 5.29 mm (0.88 %) to -0.42 mm (-0.07 %) and from 19.05 mm to 9.03 mm, respectively. It indicates that the calibration greatly improves the representation of discharge over Svartavatn. The NSE values are 0.77, 0.80 and 0.76 from Bulken, Kinne and Myrkdalsvatn, which are satisfactory although they are slightly lower than the NSE of 0.86 from Svartavatn. The infiltration parameters (refdk and refkdt) and Manning routing coefficients (mn01) are calibrated by PEST auto-calibration approach to be 3.82 E-6, 0.63 and 0.18, respectively. The correlation coefficient values of mn01 and refdk, mn01 and refkdt, refdk and refkdt are -0.23, -0.16 and 0.90, respectively. We can see that there is a high correlation between two infiltration parameters of refdk and refkdt. Figure 3 shows the daily observed discharge (black line) and simulated WRF-Hydro discharge from four study basins using various refkdt values for the extreme event during 23 - 31 October, 2014. It can be seen that WRF-Hydro is sensitive to the infiltration parameter refkdt and the uncertainty of peak flow is related to the parameters' uncertainties. The peak discharge decreases from 717, 309, 83 and 102 m³/s with REFKDT of 0.2 to 698, 217, 81 and 85 m³/s with REFKDT of 2.0 at Bulken, Kinne, Svartavatn and Myrkdalsvatn basin, respectively. An increase of refkdt in WRF-Hydro modelling leads to a decrease in peak flow, while a slower recession limb in the hydrograph. The daily observed discharge and simulated discharge based on the calibrated parameter set and the non-calibrated parameter set from the four study catchments are plotted in Figure 4. The hydrographs show that the calibrated runs capture the peak

timing and magnitude well in all four catchments and that calibration markedly improves these features. The water balance of the four study catchments is shown in Table 6, highlighting that the discharge at the four study catchments is driven by intense rainfall, and the impact of evapotranspiration (ET) and the changes in snow depth water equivalent and soil moisture are minor. ET is small for all the catchments. This is due to the low temperatures at the end of October in western Norway, which lies very close to the Arctic Circle and is dominated by mountainous terrain (Engeland et al., 2004).

4.2 Benchmark evaluation

Furthermore, the benchmark model efficiencies are also shown in Table 5. Daily precipitation, temperature and potential evaopotranspiration from WRF are used as input to the HBV light model in order to calculate benchmarks. For the upper benchmark (i.e. using calibrated parameters), the calibrated HBV model efficiency ($R_{upper)}$) of Svartavatn basin is 0.80. For the lower benchmarks, the HBV model efficiency is 0.43 when calculated from random parameters ($R_{lower/random}$) and is 0.67 when calculated from regionalized parameter sets based on three nearby catchments ($R_{lower/regional}$). Regarding the Bias and RMSE values, they are -0.42 mm (-0.07 %) and 9.03 mm from calibrated WRF-Hydro, 2.52 mm (0.42 %) and 11.3 mm from the upper benchmark, and 7.65 mm (1.27 %) / 2.95 mm (0.49 %) and 18.43 mm / 14.13 mm from two lower benchmarks ($R_{lower/random}$ / $R_{lower/regional}$). These results show that the calibrated WRF-Hydro model NSE of 0.86 is well above the upper benchmark (0.80). Besides, the calibrated WRF-Hydro has both less bias and smaller RMSE than the upper benchmark. Despite this encouraging result some care must be taken in the interpretation due to uncertainty in the input data for the HBV simulation caused by a lack of long-term averaged monthly meteorological forcing.

4.3 Precipitation evaluation

The accumulated precipitation from 23 October at 06 UTC to 31 October at 06 UTC is shown in Figure 5. The CTRL simulation (see Table 3) is shown in colored contours and the observed values in colored squares (circles) for the HOBO rain gauge (meteorological) observational network. The observed precipitation amounts correspond well with the model simulation at the majority of the stations. The spatial variability is large in the complex terrain with several areas receiving close to 500mm precipitation and some areas less than 100mm during the week.

The temporal evolution of simulated precipitation at monitoring locations is shown in Figure 6. Observational stations are depicted in Figure 6a and the simulated precipitation interpolated from the four nearest grid points in the CTRL simulation is shown in Figure 6b. Daily precipitation values are shown as diamonds, whereas hourly values are shown as continuous lines. The temporal evolution is generally well reproduced by the simulation, as well as the timing of the precipitating periods.

4.4 Sensitivity to spinup time

Five different spinup times are investigated in order to analyze the sensitivity of precipitation and discharge to the initial conditions (see list in Table 3). The same calibrated parameter set is used for all the spinup experiments.

During this event the western coast of Norway was exposed to a considerable amount of precipitation within a 4-day period. Further, the soil was already saturated after a wet October. This can be seen in Figure 7 which shows (a) the catchment averaged total soil moisture of the four catchments in the CTRL experiment (26d-spinup), (b) the averaged total soil moisture on 24 October, (c) the difference of soil moisture between the 1d-spinup and CTRL, and (d) the difference of soil moisture between the 12d-spinup and CTRL. Figure 7 indicates a sensitivity of soil moisture to spinup time although the differences are fairly small (-10 to 10 mm, which is around ± 1 %). The difference between experiments is more clearly highlighted in Figure 8, which shows the evolution of basin averaged soil moisture during the period of 23 – 31 October. The soil moisture on the first day clearly differs between spinup times in all catchments. More specifically, the soil moisture on 23 October increases with the spinup time length, which indicates that runs with short spinup have a much drier soil that can absorb additional precipitation during the initial phase of the event (i.e., 23 – 25 October). In general, the soil becomes slightly wetter with increased spinup time. 2 – 3 days after initialization, the soil is saturated no matter the spinup time. This is likely due to the relatively shallow soil depth in the mountainous region of Southern Norway.

In addition, we evaluate the temporal evolution of the precipitation for the spinup experiments. Figure 9 shows the accumulated precipitation interpolated from the four nearest grid points to the following four rain gauge stations, Øvstedal, Myrkdalen, Mjølfjell and Vossevangen. These are the official meteorological observational stations located in the catchments of Svartavatn, Myrkdalen, Kinne and Bulken, respectively. The precipitation sensitivity to spinup time is low in all catchments. At Øvstedal and Mjølfjell the precipitation is reproduced well, whereas the remaining two stations are somewhat biased. The model seems to be unable to catch the finer scale phenomena completely with a 3 km grid resolution especially over a small complex catchment such as Myrkdalvatn. This could be partly due to the combination of highly complex terrain and interactions with the Sogne Fjord, just to the north, that are missing in the simulation. Table 7 provides additional evidence of this, showing the mean absolute error (MAE) in the total accumulated precipitation and discharge. The differences in MAE of precipitation between the spinup experiments are negligible; we believe this is because of the large-scale nudging in the outer domain. Besides, there is no decrease in precipitation MAE with the spinup time increase in any of the stations, except Øvstedal. The averaged MAE of precipitation at all 54 observational stations in the area is around 50 mm. This suggests that the model, even at 3 km grid spacing, struggles to fully reproduce the local-scale orographic effects in the complex terrain around Voss and Myrkdalen. A previous study in the high mountains of Asia, suggested that a sub-kilometer grid is needed for accurately estimating truly local meteorological variability (Bonekamp et al., 2018).

The temporal evolution of streamflow over the four catchments is shown in Figure 10, which shows the daily hydrograph of discharge for the four catchments with different spinup experiments. We want to keep this flood event completely in our spinup

experiment, so we evaluate the period of 23 - 31 October. From the results, we can see that the precipitation amount and timing do not differ significantly between spinup times in any of the catchments. However, the discharge at the pre-flood phase, which is 23–24 October, is more sensitive to the spinup time. For Svartavatn this sensitive phase even extends to the 26th. This is

355   because the initial condition of soil moisture affects the overland flow that dominates the discharge of this catchment (see Figure 8). The pre-flood discharge moves closer to the observed discharge when we increase the spinup time from 1 day to 26 days, which confirms the soil moisture feedbacks from different spinup time experiments in Figure 8. In general, the peak flows are overestimated compared with the observations, except in the Svartavatn catchment. This is because we only calibrated the model in Svartavatn, and then used this calibrated parameter set in the simulation for the other three catchments,

360   which perforce have poorer performance than the Svartavatn catchment.

4.5 Sensitivity to prescribed snow cover

The dynamical modelling experiments of different snow conditions are based on the WRF-Hydro simulation with a 26-day spinup, which is labelled as the control (CTRL) in the snow experiments (Table 2). In these experiments, the snow depth and the water equivalent snow depth are changed in the 25th of October restart file and the simulations are restarted in offline

365   mode. An overview of the snow experiments is presented in Table 3. To summarize: one set of experiments tests the sensitivity to varying snow depths while the other set of experiments tests the sensitivity to snow elevation. The temporal evolution of catchment averaged SWE is shown in Figure 11 and Figure 12. A decline throughout the simulation period is shown for all catchments. The experiments with the 0.5 m, 1 m and 2 m snow depths have similar snowmelt behavior during 25 - 28 October. The snowmelt stops in all the snow experiments after 29 October, because of a drop in both rainfall and temperature (below

370   273 K), which can be seen in Figure 11. More detailed information of the total snowmelt during 25 - 31 October under the different snow experiments is given in Table 8. From the table, we can see that, except for the 0.1m snow depth experiment where the added snow quickly melts away, the results from other three prescribed snow depth experiments (0.5, 1.0 and 2.0 m) are fairly similar, where the total water equivalent snow melt are 14 -16 cm in Svartavatn, 11-12 cm in Myrkdalsvatn, 11-12 cm in Kinne and 12-13 cm in Bulken. This is because the limit of melting snow is controlled by the temperature in Noah

375   LSM and a maximum of around 0.5 m snow will be melted away in this case. For the snow elevation experiment where 1 m of snow were added above the given ground elevation, the response is a result of the elevation of the catchments. In Kinne and Myrkdalsvatn there is little variation, around 8-11 cm of SWE is melting. Their average catchment height is so high that there is only a small difference in total SWE between the experiments, leading to similar response. For Svartvatn and Bulken the situation is different. The total SWE in the catchments vary between the experiments because of the lower catchment elevation,

380   hence the resultant SWE melting varies between 2-16 cm for Svartvatn and 4-12 cm for Bulken.

Figure 13 shows the hydrograph of hourly discharge for the experiments where snow depth is modified uniformly over the entire area. From both Figures 9 and 12, we can see that the difference in melt between the shallowest and deepest snow depths

is less than 2 cm, which suggests that snow depths beyond 0.5 m did not contribute markedly to more discharge. The main contribution from the additional snow is to enhance the peak discharge in all the catchments. Also, the contribution from melting snow is mostly confined to precipitating periods, which also coincides with higher temperatures. The fact that snow depths above 0.5 m have little impact suggests that rain-on-snow can melt at most 0.5 m snow under this experiment design. The snowmelt discharge decreases after 29 October, which is preceded by a drop in both rainfall and surface temperature (below 273 k). It is worthwhile to recall that the Noah LSM has a simple bulk snow-soil-canopy layer model. Previous studies have noted that there was a positive bias in snow surface energy in the Noah LSM, which resulted in an underestimation of snow water equivalent (SWE), and led to a reduced snow pack during winter and earlier snowmelt in spring (Jin and Miller, 2007; Jin and Wen, 2012; Niu et al., 2011). In our experiments this positive bias in snow surface energy in the Noah LSM together with energy from intense rainfall are probably used to melt snow directly, so that the warm snowpack does not retain any liquid water which can refreeze during the day before it runs away. In this case, we might see unrealistically high snowmelt and the snowpack would not act like a sponge and retain part of the rainfall. All in all, the snow experiments show that intense precipitation coinciding with higher temperature can result in up to 0.5 m of snow melt, which contributes to the peak flow. However, more work needs to be done in the future with more sophisticated multi-layer snow models to confirm the behaviour observed in these idealized experiments.

The effects of varying snow cover by altitude on daily discharge are shown in Figure 14. Here, we perform snow experiments where 1m of snow is imposed above certain elevations, i.e., 400, 600, and 800 m. Those prescribed snow covers are applied in the restart file on 25 October 2014, which is from the 26-day spinup experiment with calibrated parameter set. From Figure 14 we can see that (a) there are increases in snow-melt runoff with the elevation decreases from 800 m to 0 m; (b) the differences of snow-melt runoff among different experiments vary in different catchments, for example, the snow-melt runoff from 0 m and 800 m experiments show a large difference in catchment Svartavatn, while not much difference can be seen in catchment Kinne and Myrkdalsvatn. This is because varying the prescribed snow cover by elevation has a greater influence on the lower catchments, i.e., Svartavatn (with 61 % of the area below 800 m), than the higher catchments such as Kinne and Myrkdalsvatn (with 69 % and 71 % above 800 m, respectively). A more detailed quantitative estimate of the total water equivalent snow depth change during 25-31 October under different prescribed snow cover experiments is given in Table 8. It confirms the results in Figure 11 and Figure 12, with the first half-meter of the snowpack contributing the to the snowmelt. In addition, there is a greater SWE decrease in the lower elevation catchment Svartevatn (i.e., -0.16 m in the added 1 m snow experiment) than Kinne and Myrkdalsvatn catchments (i.e., -0.11 m in the added 1 m snow experiment), which are dominated by higher elevations.

5   Discussion

Precipitation pattern in Norway varies spatially and is highly affected by the complex topography. More specifically, there is a strong west-east gradient of precipitation, with decreasing amounts as we move eastwards across the mountain range (Dyrrdal, A.V., 2015). To represent the interaction between the atmosphere and the complex terrain realistically there is a need for high spatial resolution in models. For example, for the episode under investigation here Pontoppidan et al. (2017) showed that a 3 km grid scaling represented the precipitation distribution better than an equivalent simulation with 9 km grid spacing, which lacked the observed spatial variability and was unable to show dynamical features like gravity waves. Furthermore, a recent study by Magnusson et al. (2019) found that the grid resolution is important in energy-balanced snow models and the scale error increases with subgrid topographic variability. They also suggested that for snow models the best is to run at the highest possible resolution and any upscaling can bring large regional errors because of model nonlinearities. The results from our study confirmed that convection-permitting simulations fairly fit the requirements of hydrological processes determining flood events in western Norway and address them in a realistic manner.

Most of the precipitation in Norway is frontal, caused by large-scale cyclone activity in the North Atlantic (Heikkilä et al., 2011). In the West coast region, extreme precipitation occurs in autumn and winter, which is dominated by orography and frontal systems (Dyrrdal, A.V., 2015). In Eastern Norway, where the mountain ranges are located, the annual precipitation is less than the west but with the highest amounts occurring near the steepest surface slopes in winter and fall (Andersen, 1972). In Southeast Norway, however, intense precipitation is dominated by convective precipitation in summer. Norway, despite its high latitude, has a diverse range of climates including northern Arctic, central alpine and southern maritime and can exhibit an equally wide range of snow regimes (Pall et al., 2019). The role of snowmelt and rainfall is highly relevant for the seasonal flood regimes (Barnett et al., 2005; Vormoor et al., 2016). For example, south-central Norway has an alpine climate, which receives large amounts of precipitation, approximately 30% as snowfall (Saloranta 2014) and has high discharge during spring and early summer due to snowmelt. Southwestern Norway has a maritime climate and the highest precipitation occurs in fall and winter, which often results in flood events (Vormoor et al. 2016). From the results of the study, we can see that the snow feedback to river flow depends on which snow regime the region is in, i.e., little or no snow (CTRL, 0.1 m snow), a lot of snow (1.0 or 2.0 m snow), or somewhere-in-between (0.5 m snow). According to previous studies of the current climate, snow cover above 800m is present for over 200 days of the year in southern Norway (Hanssen-Bauer et al., 2015) and the observed median snow depth of Norway varies from 0.1 - 0.5 m during October - May (1957-2011) (Saloranta, 2012). In some regions of Southern Norway, the snow depth can be up to 2 - 3 m during the late winter (Andreassen and Oerlemans, 2009). Furthermore, Pall et al. (2019) has constructed a rain-on-snow climatology using a 1km gridded observation data (during 1961–1990) and found that an average monthly count of daily rain-on-snow events varies from 2 to 4 during winter-spring in Southern Norway. Under climate change impact, the snowpack distribution (both temporal and spatial) in Norway will be changing. In general, snowmelt floods will reduce in Norway, while the winter precipitation will increase, which may also

lead to larger snow storage, e.g. in mountainous areas in Eastern Norway (Hanssen-Bauer et al. 2015). Meanwhile, other
445 studies also showed that in Norway general increases in both precipitation and temperature (especially warmer winters) will intensificate the risk of rain-on-snow events in certain regions and seasons. Such events can be a major trigger of hazards, i.e., floods and landslides, in the country (Hansen et al. 2014; Pall et al., 2019). The regional pattern of increases and decreases of flood events (both frequency and magnitude) reflects the balance between the different and sometimes counteracting processes, e.g., snowpack dynamics, snowfall vs. rainfall.

6   Conclusions

In this study, we aimed to reproduce an extreme weather event in a region characterized by complex terrain. A dynamical hydrometeorological modeling system (WRF-Hydro) was employed for this purpose. A nested WRF atmospheric model, run at convection permitting scales, was used to reproduce the meteorological event and provide precipitation forcing for a distributed hydrological model over a small domain encompassing four study catchments affected by extreme flooding. 3 km
455 grid spacing was used for the WRF atmosphere and land surface while a 300 m grid spacing was used for the WRF-Hydro river routing. An auto-calibration tool was used for WRF-Hydro model calibration based on the daily discharge at Svartevatn, which is the smallest of the 4 catchments. The simulation of high-resolution precipitation and discharge were assessed based on observational data sets. Also, the sensitivity of the results to the spinup time and snow depth was investigated.

The results showed that the precipitation from the 3 km simulation generally agreed well with the rain gauges both in terms of
460 temporal evolution and spatial variability, although it underestimated the precipitation in the highly complex terrain around Myrkdalen. This underestimation could be due to a combination of the locally complex topography and the proximity to the Sognefjord only 15 km away. This large, but narrow body of water, and its many offshoots, was not well-resolved by the modeling system.

The auto-calibration greatly improved the model performance with the NSE increasing from 0.41 to 0.86, bias and RMSE
465 decreasing from 5.29 and 19.05 mm to -0.42 and 9.03 mm, respectively. The modeling system captured peak flow volumes and timing well after model calibration. Besides, WRF-Hydro runoff performance depended on some highly sensitive parameters (e.g. infiltration parameter and Manning routing coefficients).

Comparing with the benchmarks, the calibrated WRF-Hydro NSE value (0.86) was higher than the upper benchmark from the HBV light model (0.80). This might be due a lack of long-term input data (e.g., averaged monthly potential evapotranspiration).
470 The implication was that WRF-Hydro might perform as well or even better than a simpler conceptual hydrological model, especially for ungauged basins or observation scarce regions.

The precipitation simulation was not overly sensitive to spinup time. We found that mean absolute errors of precipitation were very similar given different spinup times. This could be due, in part, to the decision to nudge the atmospheric flow to match

that large-scale reanalysis. Discharge simulations were slightly more sensitive to the spinup time due to the impact of soil moisture, especially during the pre-peak phase. We found a spinup time of 26 days give the lowest MAE of precipitation and discharge compared to the other smaller periods.

SWE melt during 25-31 October was consistently around 10 –16cm for the uniform snow depth experiments (0.5 - 2 m). The results also showed that melting snow contributes most to discharge during the rainy periods and the peak flow periods. This indicated that snow cover intensified the extreme discharge instead of acting as sponge in this study, which suggested that future rain-on-snow events might potentially result in higher flood risk. However, more sophisticated snow models and targeted experiments should be conducted to confirm this speculation.

Our results increased confidence in the performance of WRF-Hydro for simulating extreme hydrometeorological events over complex terrain. Further, they demonstrated the importance of model calibration and reasonably accurate land surface initial conditions for simulating discharge, especially for peak flow. The snow experiments suggested that rain-on-snow events under warmer conditions might contribute to an increase in flood magnitudes in Norway, due to projected increases in extreme precipitation (Lawrence, 2016). However, targeted experiments on the changing risks associated with future rain-on-snow events are needed to confirm this possibility.

Acknowledgments

This study was joint funded by the Research Council of Norway (RCN) project EVOGLAC (grant 255049), HordaKlim (grant 245403) and R3 (grant 255397), by the Bjerknes Center for Climate Research (SKD) project WACYEX and CHEX. The computer resources where available through the RCN's program for supercomputing (NOTUR/NORSTORE); projects NN9280K, NN9478K and NS9001K. All WRF-Hydro simulation data in this paper are available from the authors upon request (luli@norceresearch.no).

Author contributions. LL contributed to most of the modelling, analysis, writing and revising of the paper. MP contributed to collecting HOBO data, analysis and assisted with writing and reviewing the paper. SS contributed to the reviewing the paper. AS contributed to the model calibration by PEST and reviewing the paper.

Competing interests. The authors declare that they have no conflict of interest.

References

Andersen, P., 1972. The distribution of monthøy precipitation in southern Norway in relation to prevailing H.Johansen weather types. Yearbook for Univ. og Bergen Mat., Naturv. Series, No.1.

Andreassen, L.M. and Oerlemans, J., 2009. Modelling long-term summer and winter balances and the climate sensitivity of storbreen, norway. Geografiska Annaler: Series A, Physical Geography, 91(4), pp.233-251.

505 Arnault, J., Wagner, S., Rummler, T., Fersch, B., Bliefernicht, J., Andresen, S., & Kunstmann, H. (2016). Role of Runoff–Infiltration Partitioning and Resolved Overland Flow on Land–Atmosphere Feedbacks: A Case Study with the WRF-Hydro Coupled Modeling System for West Africa. Journal of Hydrometeorology, 17(5), 1489–1516. https://doi.org/10.1175/JHM-D-15-0089.1

Avolio, E., Cavalcanti, O., Furnari, L., Senatore, A., and Mendicino, G. (2019). Brief communication: Preliminary hydro-510 meteorological analysis of the flash flood of 20 August 2018 on Raganello Gorge, Southern Italy, Nat. Hazards Earth Syst. Sci., 19, 1619-1627. https://doi.org/10.5194/nhess-19-1619-2019.

Barnett, T.P., Adam, J.C. and Lettenmaier, D.P., 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. Nature, 438(7066), p.303.

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the 515 United States. Geophysical Research Letters, 43(9), 4382–4390. https://doi.org/10.1002/2016GL068070

Bonekamp, P.N.J., Collier, E. and Immerzeel, W.W., 2018. The Impact of Spatial Resolution, Land Use, and Spinup Time on Resolving Spatial Precipitation Patterns in the Himalayas. Journal of Hydrometeorology, 19(10), pp.1565-1581.

Collier, E., Mölg, T., Maussion, F., Scherer, D., Mayer, C., & Bush, A. B. G. (2013). High-resolution interactive modelling of the mountain glacier–atmosphere interface: an application over the Karakoram. The Cryosphere Discussions, 7(1), 103-144.

520 Dannevig et al., 2016, Naturfareprosjektet Oktoberflaumen på Vestlandet i 2014, rapport 2016-36, in Norwegian.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quart. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/ 10.1002/qj.828.

Dyrrdal, A. V., Isaksen, K., Hygen, H. O., & Meyer, N. K. (2012). Changes in meteorological variables that can trigger natural hazards in Norway. Climate Research, 55(2), 153–165. https://doi.org/10.3354/cr01125

525 Dyrrdal, A.V., 2015. Estimating extreme precipitation on different spatial and temporal scales in Norway.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. J. Geophys. Res., 108, 8851, doi:10.1029/2002JD003296.

El-Samra, R., Bou-Zeid, E., & El-Fadel, M. (2018). What model resolution is required in climatological downscaling over 530 complex terrain?. Atmospheric research, 203, 68-82.

Freudiger, D., Kohn, I., Stahl, K., & Weiler, M. (2014). Large-scale analysis of changing frequencies of rain-on-snow events with flood-generation potential. Hydrology and Earth System Sciences, 18(7), 2695–2709. https://doi.org/10.5194/hess-18-2695-2014

Engeland, K., Skaugen, T.E., Haugen, J.E., Beldring, S. and Førland, E.J., 2004. Comparison of evaporation estimated by the
535 HIRHAM and GWB models for present climate and climate change scenarios. Norwegian Meteorological Institute, met (No. 17). no Report.

Givati, A., Gochis, D., Rummler, T. and Kunstmann, H., 2016. Comparing one-way and two-way coupled hydrometeorological forecasting systems for flood forecasting in the Mediterranean region. Hydrology, 3(2), p.19.

Gochis, D.J., W. Yu, D.N. Yates, 2015: The WRF-Hydro model technical description and user's guide, version 3.0. NCAR
540 Technical Document. 120 pages. Available at: WRF-Hydro 3.0 User Guide.

Gochis, D.J., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, J. McCreight, J. Mills, A. RafieeiNasab, L. Read, K. Sampson, D. Yates, W. Yu, (2018). The WRF-Hydro modeling system technical description, (Version 5.0). NCAR Technical Note. 107 pages. Available online at https://ral.ucar.edu/sites/default/files/public/WRF-HydroV5TechnicalDesc. Source Code DOI:10.5065/D6J38RBJ

545 Hansen, B.B., Isaksen, K., Benestad, R.E., Kohler, J., Pedersen, Å.Ø., Loe, L.E., Coulson, S.J., Larsen, J.O. and Varpe, Ø., 2014. Warmer and wetter winters: characteristics and implications of an extreme weather event in the High Arctic. Environmental Research Letters, 9(11), p.114021.

Hansen, B.B., Isaksen, K., Benestad, R.E., Kohler, J., Pedersen, Å.Ø., Loe, L.E., Coulson, S.J., Larsen, J.O. and Varpe, Ø., 2014. Warmer and wetter winters: characteristics and implications of an extreme weather event in the High Arctic.
550 Environmental Research Letters, 9(11), p.114021.

Hanssen-Bauer, I., and Coauthors, 2015: Klima i Norge 2100. Norwegian Centre for Climate Services Rep. 2/2015, 204 pp.

Hanssen–Bauer, I. and Førland, E., 2000. Temperature and precipitation variations in Norway 1900–1994 and their links to atmospheric circulation. International Journal of Climatology: A Journal of the Royal Meteorological Society, 20(14), pp.1693-1708.

555 Hanssen-Bauer, I., Førland, E. J., Haddeland, I., Hisdal, H., Lawrence, D., Mayer, S., Ådlandsvik, B. (2017). Climate in Norway 2100.

Heikkilä, U., Sandvik, A. and Sorteberg, A., 2011. Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. Climate dynamics, 37(7-8), pp.1551-1564.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by
560 long–lived greenhouse gases: Calculations with the AER radiative transfer models. J. Geophys. Res., 113, D13103. doi:10.1029/2008JD009944.

Jin, J. and Miller, N.L., 2007. Analysis of the impact of snow on daily weather variability in mountainous regions using MM5. Journal of Hydrometeorology, 8(2), pp.245-258.

Jin, J. and Wen, L., 2012. Evaluation of snowmelt simulation in the Weather Research and Forecasting model. Journal of
565 Geophysical Research: Atmospheres, 117(D10).

Julien, P., B. Saghafian, and F. Ogden, 1995: Raster-based hydrological modeling of spatially-varied surface runoff. Water Resour. Bull., 31 (3), 523-536.

Hong, Song–You, Yign Noh, Jimy Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Wea. Rev., 134, 2318–2341. doi:10.1175/MWR3199.1

570   Kendon, E. J., Ban, N., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., ... & Wilkinson, J. M. (2017). Do convection-permitting regional climate models improve projections of future precipitation change?. Bulletin of the American Meteorological Society, 98(1), 79-93.

Kerandi, N., Arnault, J., Laux, P., Wagner, S., Kitheka, J., & Kunstmann, H. (2018). Joint atmospheric-terrestrial water balances for East Africa: a WRF-Hydro case study for the upper Tana River basin. Theoretical and Applied Climatology,
575   131(3–4), 1337–1355. https://doi.org/10.1007/s00704-017-2050-8

Langsholt, E., Roald, L. A., Holmqvist, E., Fleig, A., 2015, Flommen på Vestlandet oktober 2014, NVE rapport 2015-11, in Norwegian.

Lawrence, D., 2016. Klimaendring og framtidige flommer i Norge. NVE Rapport nr. 81-2016.

Mlawer, Eli. J., Steven. J. Taubman, Patrick. D. Brown, M. J. Iacono, and S. A. Clough,1997: Radiative transfer for
580   inhomogeneous atmospheres: RRTM, a validated correlated–k model for the longwave. J. Geophys. Res., 102, 16663–16682. doi:10.1029/97JD00237.

Li, L., Gochis, D. J., Sobolowski, S., & Mesquita, M. D. S. (2017). Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model. Journal of Geophysical Research, 122(9), 4786–4807. https://doi.org/10.1002/2016JD026279

585   Lin, P., Rajib, M.A., Yang, Z.L., Somos–Valenzuela, M., Merwade, V., Maidment, D.R., Wang, Y. and Chen, L., 2018. Spatiotemporal Evaluation of Simulated Evapotranspiration and Streamflow over Texas Using the WRF–Hydro–RAPID Modeling Framework. JAWRA Journal of the American Water Resources Association, 54(1), pp.40-54. https://doi.org/10.1111/1752-1688.12585

Livneh, B., Xia, Y., Mitchell, K.E., Ek, M.B. and Lettenmaier, D.P., 2010. Noah LSM snow model diagnostics and
590   enhancements. Journal of Hydrometeorology, 11(3), pp.721-738.

Magnusson, J., Eisner, S., Huang, S., Lussana, C., Mazzotti, G., Essery, R., Saloranta, T. and Beldring, S., Influence of spatial resolution on snow cover dynamics for a coastal and mountainous region at high latitudes (Norway). Water Resources Research. 55, 5612–5630.

Marks, D., Kimball, J., Tingey, D., & Link, T. (1998). The sensitivity of snowmelt processes to climate conditions and forest
595   cover during rain-on-snow: a case study of the 1996 Pacific Northwest flood. Hydrological Processes, 12(10–11), 1569–1587. https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11<1569::AID-HYP682>3.0.CO;2-L

Marks, D., Link, T., Winstral, A., & Garen, D. (2001). Simulating snowmelt processes during rain-on-snow over a semi-arid mountain basin. Annals of Glaciology, 32, 195–202. https://doi.org/10.3189/172756401781819751

Maussion, F., Scherer, D., Mölg, T., Collier, E., Curio, J., & Finkelnburg, R., 2014. Precipitation Seasonality and Variability over the Tibetan Plateau as Resolved by the High Asia Reanalysis. Journal of Climate, 27(5), 1910-1927.Musselman, K.N., Lehner, F., Ikeda, K., Clark, M.P., Prein, A.F., Liu, C., Barlage, M. and Rasmussen, R., 2018. Projected increases and shifts in rain-on-snow flood risk over western North America. Nature Climate Change, 8(9), p.808.

Mitchell, K. E., and Coauthors, 2001: The Community Noah Land Surface Model (LSM)—user's guide (v2.2). Available online at http://www.emc.ncep.noaa.gov/mmb/gcp/noahlsm/README_ 2.2.htm.

Naabil, E., Lamptey, B.L., Arnault, J., Olufayo, A. and Kunstmann, H., 2017. Water resources management using the WRF-Hydro modelling system: Case-study of the Tono dam in West Africa. Journal of Hydrology: Regional Studies, 12, pp.196-209.

Nash, J. E. & Sutcliffe, J. V. 1970. River flow forecasting through conceptual models – Part 1 – A discussion of principles. Journal of Hydrology 10 (3), 282–290.

Niu, G.-Y., and Coauthors, 2011: The community Noah land sur- face model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measure- ments. J. Geophys. Res., 116, D12109, https://doi.org/10.1029/ 2010JD015139.

Pall, P., Tallaksen, L.M. and Stordal, F., 2019. A climatology of rain-on-snow events for Norway. Journal of Climate, 32(20), pp.6995-7016.

Pontoppidan, M., Reuder, J., Mayer, S., & Kolstad, E. W. (2017). Downscaling an intense precipitation event in complex terrain: the importance of high grid resolution. Tellus A: Dynamic Meteorology and Oceanography, 69(1), 1271561. https://doi.org/10.1080/16000870.2016.1271561

Poschlod, B., Hodnebrog, Ø., Wood, R. R., Alterskjær, K., Ludwig, R., Myhre, G., & Sillmann, J. (2018). Comparison and Evaluation of Statistical Rainfall Disaggregation and High-Resolution Dynamical Downscaling over Complex Terrain. Journal of Hydrometeorology, 19(12), 1973-1982.

Prein, A.F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F. and Brisson, E., 2015. A review on regional convection–permitting climate modeling: Demonstrations, prospects, and challenges. Reviews of Geophysics, 53(2), pp.323-361.

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., ... & Vautard, R. (2016). Precipitation in the EURO-CORDEX 0.11◦ and 0.44◦ simulations: high resolution, high benefits? Climate dynamics, 46(1-2), 383-412.

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. Nature Climate Change, 7(1), 48.

Radu R, De ´que ´ M, Somot S (2008) Spectral nudging in a spectral regional climate model. Tellus 60A:898–910.

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., ... & Gutmann, E., 2011. High-resolution coupled climate

630   runoff simulations of seasonal snowfall over Colorado: a process study of current and warmer climate. Journal of Climate, 24(12), 3015-3048. doi: http://dx.doi.org/10.1175/2010JCLI3985.1

Rasmussen, R.M., K. Ikeda, C. Liu, D.J. Gochis, M. Clark, A. Dai, E. Gutmann, J. Dudhia, F. Chen, M.J. Barlage, D. Yates, and G. Zhang, 2014: Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations. Journal of Hydrometeorology, 15, 1091-1116, DOI: 10.1175/JHM-D-13-0118.1.

635   Reuder, J., Fagerlid, G. O., Barstad, I., Sandvik, A., 2007, Stord Orographic Precipitation Experiment (STOPEX): an overview of phase I, Advances in Geosciences, 10, 17-23. https://doi.org/10.5194/adgeo-10-17-2007.

Román-Cascón, C., G. Steeneveld, C. Yagüe, M. Sastre, J. Arrillaga, and G. Maqueda, 2016: Forecasting radiation fog at climatologically contrasting sites: Evaluation of statistical methods and WRF. Quart. J. Roy. Meteor. Soc., 142, 1048– 1063, https://doi.org/10.1002/qj.2708.

640   Rummler, T. et al., 2019: Role of Lateral Terrestrial Water Flow on the Regional Water Cycle in a Complex Terrain Region: Investigation With a Fully Coupled Model System. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES Volume: 124 Issue: 2 Pages: 507-529.

Rusli, S.R., Yudianto, D. and Liu, J.T., 2015. Effects of temporal variability on HBV model calibration. *Water Science and Engineering*, *8*(4), pp.291-300.

645   Saloranta, T.M., 2012. Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model. The Cryosphere, 6(6), pp.1323-1337.

Sampson, K. and Gochis, D., WRF Hydro GIS Pre-Processing Tools, Version 5.0, Documentation, 2018.

Seibert, J. and Vis, M.J., 2012. Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*(9), pp.3315-3325.

650   Seibert, J., Vis, M.J., Lewis, E. and Meerveld, H.V., 2018. Upper and lower benchmarks in hydrological modelling. *Hydrological Processes*.

Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N., & Kunstmann, H. (2015). Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. Journal of Advances in Modeling Earth Systems, 7(4), 1693–1715. https://doi.org/10.1002/2015MS000510.

655   Smiatek, G., Kunstmann, H., & Senatore, A. (2016). EURO–CORDEX regional climate model analysis for the Greater Alpine Region: Performance and expected future change. Journal of Geophysical Research: Atmospheres, 121(13), 7710-7728.

Sorteberg, A., Lawrence, D., Dyrrdal, A. V., Mayer, S., & Engeland, K. (2018). Climatic changes in short duration extreme precipitation and rapid onset flooding - implications for design values. Retrieved from https://cms.met.no/site/2/klimaservicesenteret/rapporter-og-publikasjoner/_attachment/13537?_ts=163df95ff7b

660 Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., Van Lanen, H. A. J., Sauquet, E., … Jódar, J. (2010). Streamflow trends in Europe: Evidence from a dataset of near-natural catchments. Hydrology and Earth System Sciences, 14(12), 2367–2382. https://doi.org/10.5194/hess-14-2367-2010.

Verri, G., Pinardi, N., Gochis, D., Tribbia, J., Navarra, A., Coppini, G., & Vukicevic, T. (2017). A meteo-hydrological modelling system for the reconstruction of river runoff: The case of the Ofanto river catchment. Natural Hazards and Earth

665 System Sciences, 17(10), 1741–1761. https://doi.org/10.5194/nhess-17-1741-2017.

Vormoor, K., Lawrence, D., Heistermann, M., & Bronstert, A. (2015). Climate change impacts on the seasonality and generation processes of floods – projections and uncertainties for catchments with mixed snowmelt/rainfall regimes. Hydrology and Earth System Sciences, 19(2), 913–931. https://doi.org/10.5194/hess-19-913-2015.

Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D., & Wong, W. K. (2016). Evidence for changes in the magnitude and

670 frequency of observed rainfall vs. snowmelt driven floods in Norway. Journal of Hydrology, 538, 33–48. https://doi.org/10.1016/j.jhydrol.2016.03.066.

Wilson, D., Hisdal, H., & Lawrence, D. (2010). Has streamflow changed in the Nordic countries? - Recent trends and comparisons to hydrological projections. Journal of Hydrology, 394(3–4), 334–346. https://doi.org/10.1016/j.jhydrol.2010.09.010.

675 Xiang, T., Vivoni, E.R., Gochis, D.J. and Mascaro, G., 2017. On the diurnal cycle of surface energy fluxes in the North American monsoon region using the WRF–Hydro modeling system. Journal of Geophysical Research: Atmospheres, 122(17), pp.9024-9049.

Yucel, I., A. Onen, K. K. Yilmaz, and D. J. Gochis, 2015. Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall, J. Hydrol., 523, 49–66.

680

Table 1: The list of observed discharge stations and weather stations located in each of the catchments

Discharge station	Area (km²)	Latitude	Longitude	Nearest rain gauge
Bulken	1092.04	6.29	60.63	Vossevangen
Kinne	511.8	6.50	60.63	Mjølfjell
Svartavatn	72.3	5.90	60.65	Øvstedal
Myrkdalsvatn	158.87	6.50	60.80	Myrkdalen

Table 2: Overview of performed initialization experiments. All experiments are based on the calibrated parameter-set.

Name	Spinup time
1d-spinup	1 day
3d-spinup	3 days
5d-spinup	5 days
14d-spinup	14 days
26d-spinup*	26 days

*: 26d-spinup is used as the control (CTRL) experiment in the snow experiments.

685

Table 3: Overview of pre-existing snow cover experiments. All experiments are performed during 25 - 31 October, based on the 26 days spinup simulation with calibrated parameter-set.

Name	Added snow depth (m)	Added water equivalent snow depth* (m)	Elevation limits for adding snow** (a.s.l. m)
2m snow	2	0.67	0
1m snow	1	0.33	0
0.5m snow	0.5	0.17	0
0.1m snow	0.1	0.03	0
800m Elev	1	0.33	800
600m Elev	1	0.33	600
400m Elev	1	0.33	400

| 0m Elev | 1 | 0.33 | 0 |

* the snow density is assumed to be 300 kg/m^3
** the elevation limits for adding snow is 0 m, which means the snow is added over the whole catchment area.

690

Table 4: Area-elevation of the four catchments

>Elevation	400 m	600 m	800 m	1000 m
Bulken (%)	85	73	57	36
Kinne (%)	92	85	69	48
Svartavatn (%)	89	67	39	9
Myrkdalsvatn (%)	94	83	71	44

695   Table 5: WRF-Hydro uncalibrated, calibrated, validated results with model efficiency values for different benchmarks

	Catchment	NSE	Bias(mm)	Bias (%)	RMSE(mm)
Default (non-calibration)	Svartavatn	0.41	5.29	0.88	19.05
Calibration	Svartavatn	0.86	-0.42	-0.07	9.03
Upper benchmark	Svartavatn	0.80	2.52	0.42	11.30
Lower benchmark (random parameter values)	Svartavatn	0.43	7.65	0.94	18.43
Lower benckmark (regional parameter values)	Svartavatn	0.67	2.95	0.49	14.13
Validation	Bulken	0.77	0.99	0.24	6.64
	Kinne	0.80	-1.52	-0.41	5.80

	Myrkdalsvatn	0.76	-0.005	0	6.65

Table 6: The water balance of four study catchments during 25 - 31 October 2014 (based on calibrated parameter sets, unit: mm)

Catchments	Observed Discharge	Precipitation	Discharge	ET	Soil water	Water equivalent snow depth	Residuals
Svartavatn	315	293.6	305.7	1.5	3.2	0.3	-17.1
Bulken	243.2	230.3	236.7	2.1	-2.5	0.9	-7
Kinne	187.6	206.6	219.8	1.2	-11.7	1.2	-3.9
Myrkdalsvatn	215	211.1	208	1.9	-14.7	0.6	15.3

700

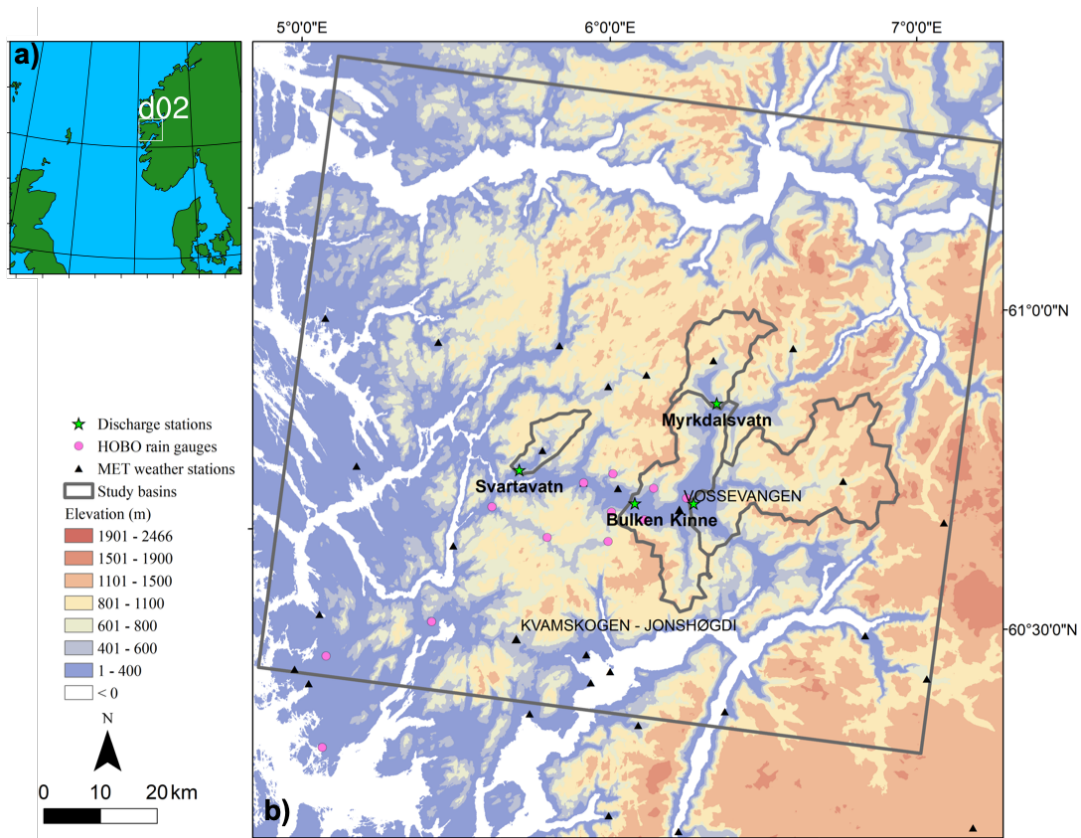
Table 7: Mean absolute error [mm] of precipitation and discharge compared with observations under impact of spinup time (23-31 October). The accumulated precipitation was interpolated from the four nearest grid points to Øvstedal, Myrkdalen, Mjølfjell, Vossevangen and (ALL) the 54 observational stations available in the area.

MAE [mm]	spinup	Øvstedal (Svartavatn)	Myrkdalen (Myrkdalsvatn)	Mjølfjell (Kinne)	Vossevangen (Bulken)	ALL
	1d	20	115	31	72	51
	3d	19	114	36	71	50
Precipitation	5d	5	113	34	69	52
	12d	5	114	31	73	51
	26d	5	125	32	69	51
	1d	108	35	9	53	29
Discharge (m3/s)	3d	80	32	13	48	20

5d	56	19	29	35	11
12d	59	17	26	31	13
26d	37	19	23	35	8

Table 8: The total water equivalent snow depth change during 25-31 October under different pre-existing snow cover experiment in four study catchments.

Experiments	Svartavatn	Myrkdalsvatn	Kinne	Bulken
CTRL	0.0004	0.0006	0.0014	0.0012
2m snow	-0.16	-0.12	-0.12	-0.13
1m snow*	-0.16	-0.11	-0.11	-0.12
0.5m snow	-0.14	-0.11	-0.11	-0.11
0.1m snow	-0.02	-0.02	-0.02	-0.02
800m Elev	-0.02	-0.08	-0.08	-0.04
600m Elev	-0.09	-0.09	-0.10	-0.09
400m Elev	-0.12	-0.11	-0.11	-0.11
0m Elev*	-0.16	-0.11	-0.11	-0.12

*: 1m snow and 0m Elev are the same experiment.

715 **Figure 1: The study catchments locate at western Norway. (a) Outer domain (resolution 9 km), borders of the inner domain (d02) are shown with a white frame; (b) zoom on the inner domain (resolution 3 km); topography and four study catchments boundaries are shown with details about ground-based observation stations, including discharge stations, HOBO rain gauges and MET weather stations (from met.no), for model calibration and further performance evaluation.**

720

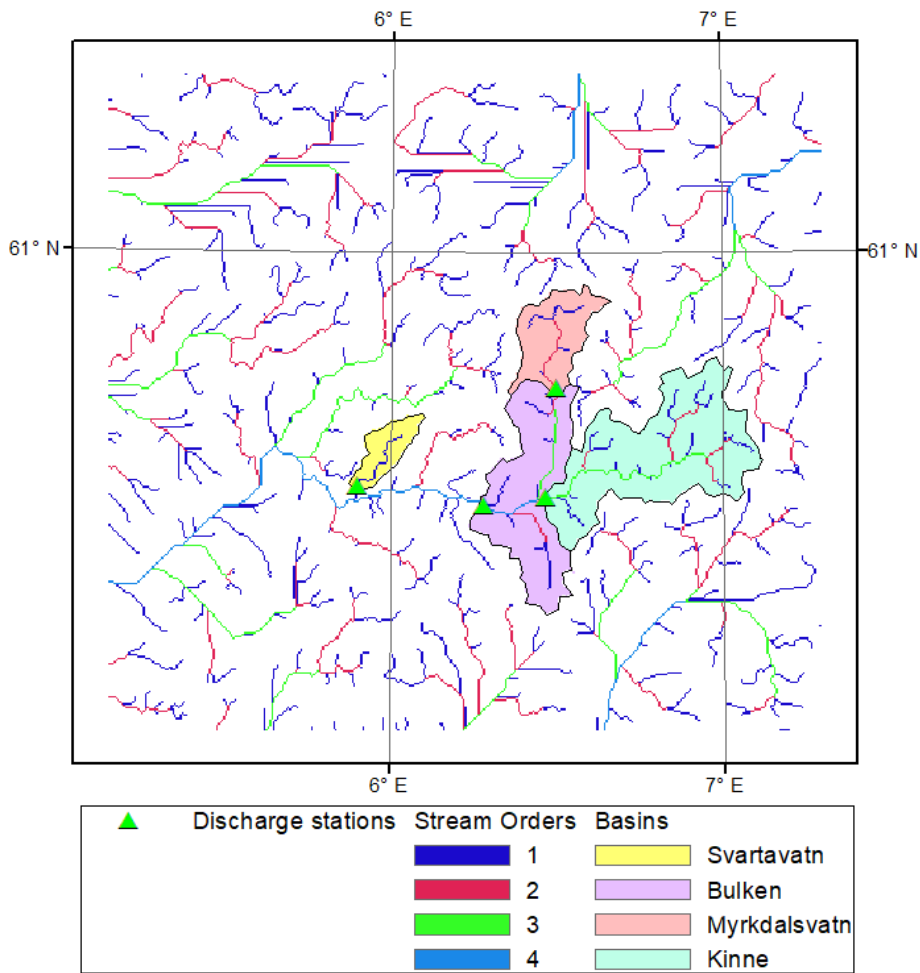



Figure 2: Map of physiographic grid for routing processes with the four study catchments (i.e., Svartavatn, Bulken, Myrkdalsvatn and Kinne), the discharge stations and stream orders in the inner domain from hydrological modelling in WRF-Hydro model system. It is worth to note that the Bulken is at the downstream of Kinne and Myrkdalsvatn, which means the drainage basin of Bulken catchment comprises the catchment of Kinne (pink) and Myrkdalsvatn (light green) in the study.

725

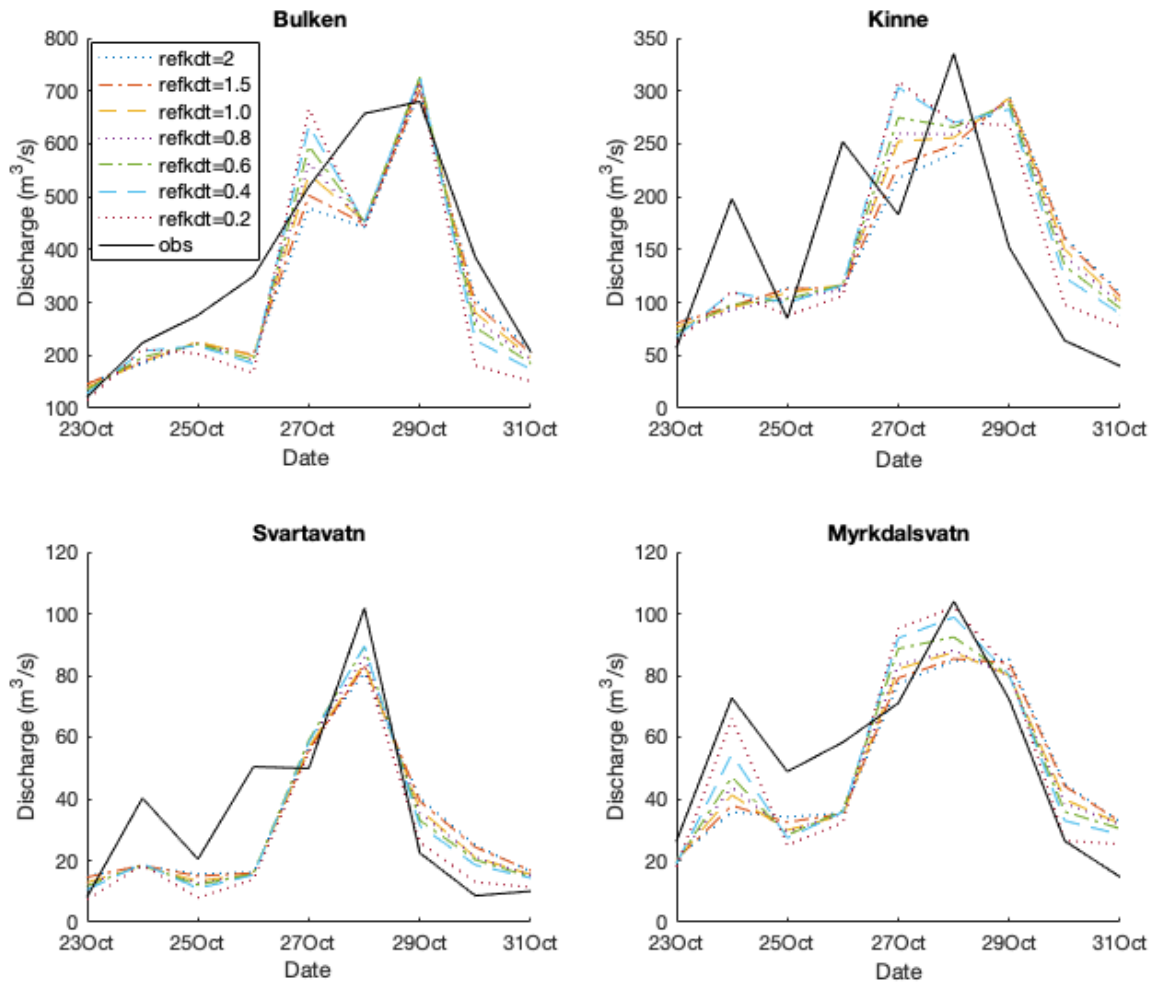



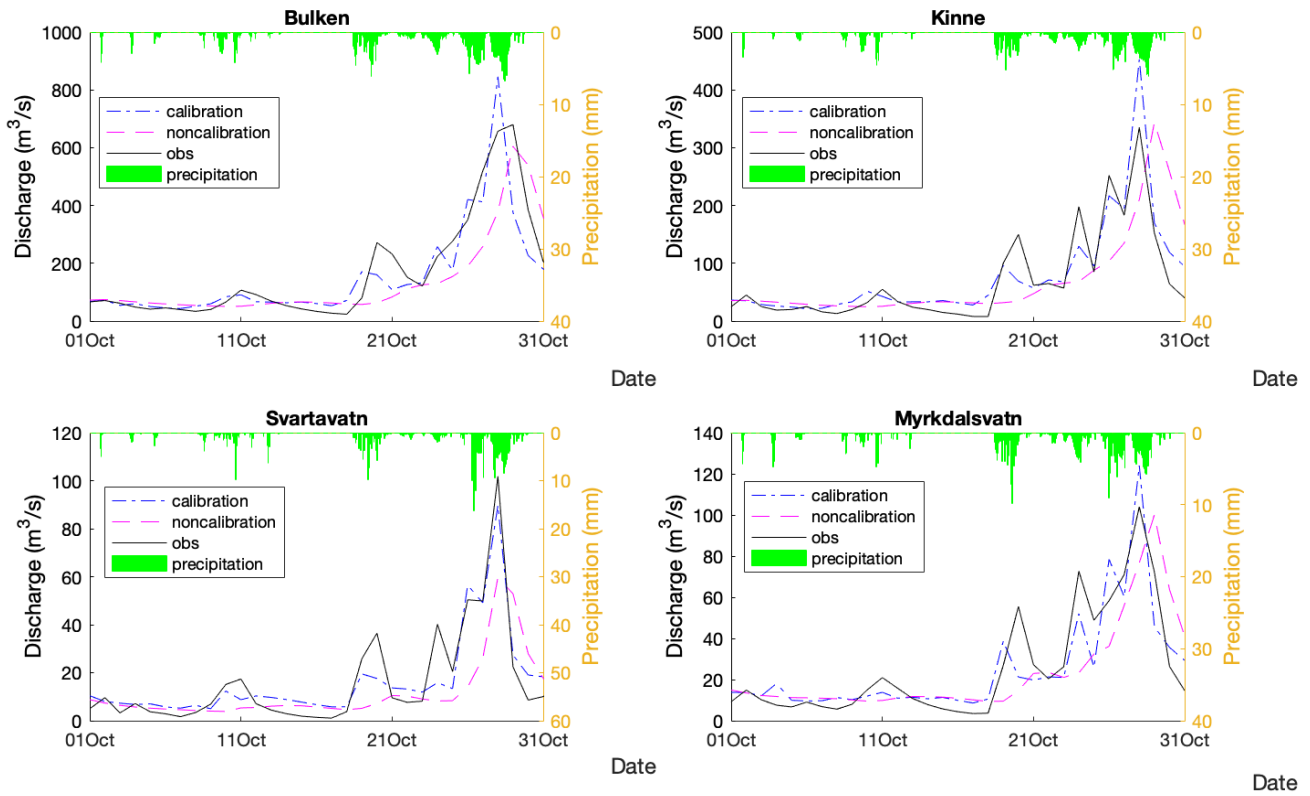
Figure 3: Hydrographs of daily observed discharge (obs) and simulated WRF-Hydro discharge from four study basins using various refkdt values for the extreme event during 23 - 31 October, 2014.

730

Figure 4: Hydrographs of daily observed discharge (obs) and simulated discharge with calibrated parameters (calibration) and non-calibrated parameters (noncalibration) with precipitation during 1 - 31 October, 2014.

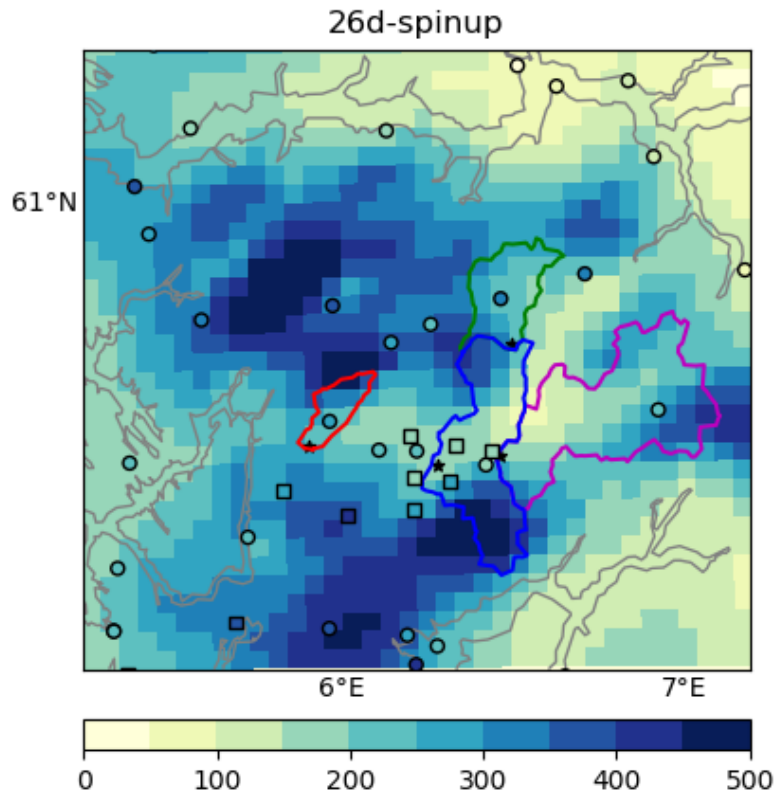



Figure 5: Accumulated precipitation (mm) during 23 - 31 October from the 26 day spin up CTRL simulation. The catchments are contoured in colours corresponding to Figure 9. Squares and circles are observational values from the hobo network and the meteorological network respectively.

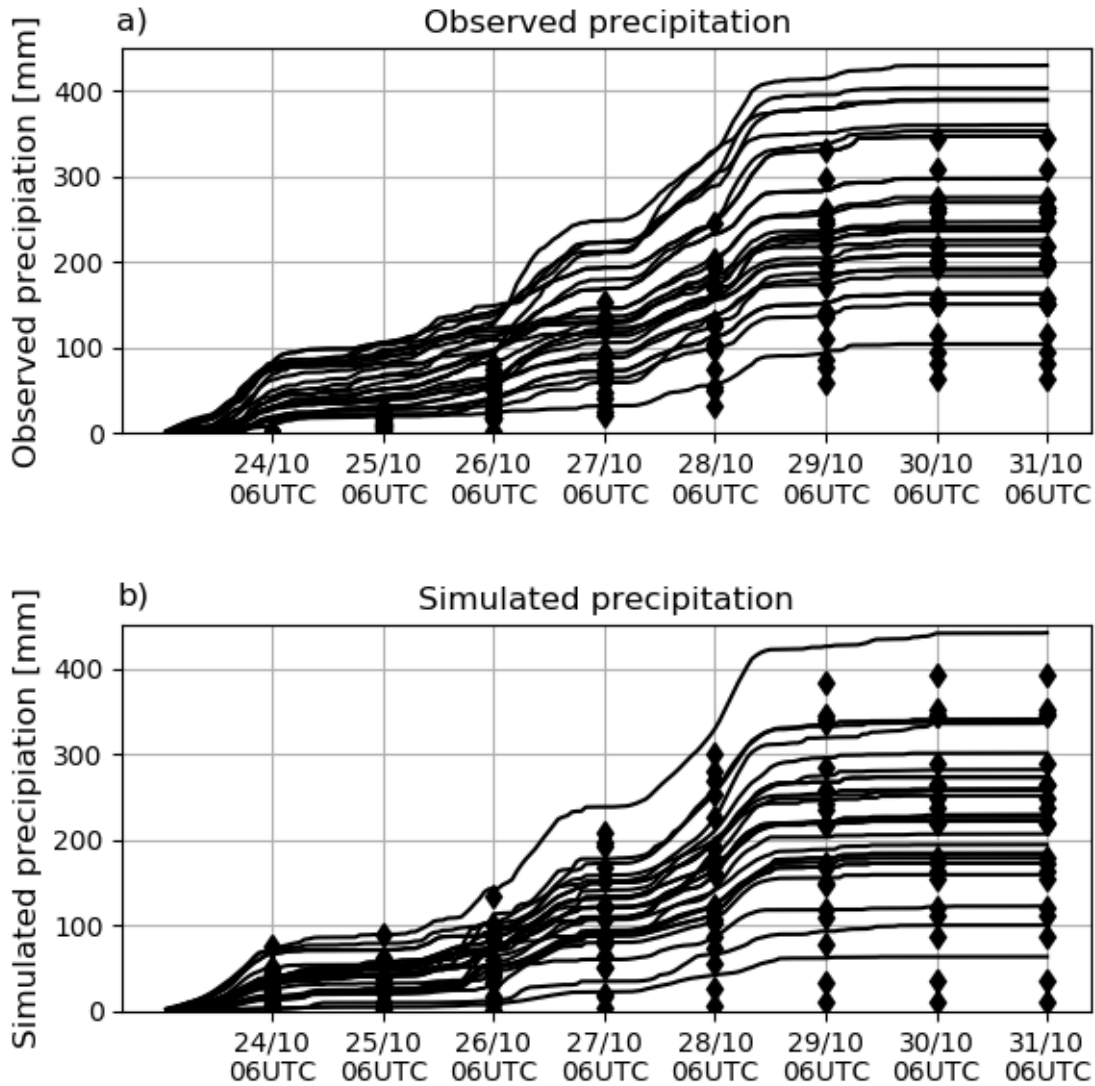



Figure 6: Accumulated precipitation from 23 October at 6 UTC to 31 October at 6 UTC a) at observation stations in the area and b) interpolated from the four nearest grid points to the equivalent station position in the 26d spin up simulation. The lines represent stations with hourly precipitation measurements whereas the diamonds represent stations with daily precipitation values. The same notation is used in b), though the model output enables higher temporal resolution at the daily station positions.

740

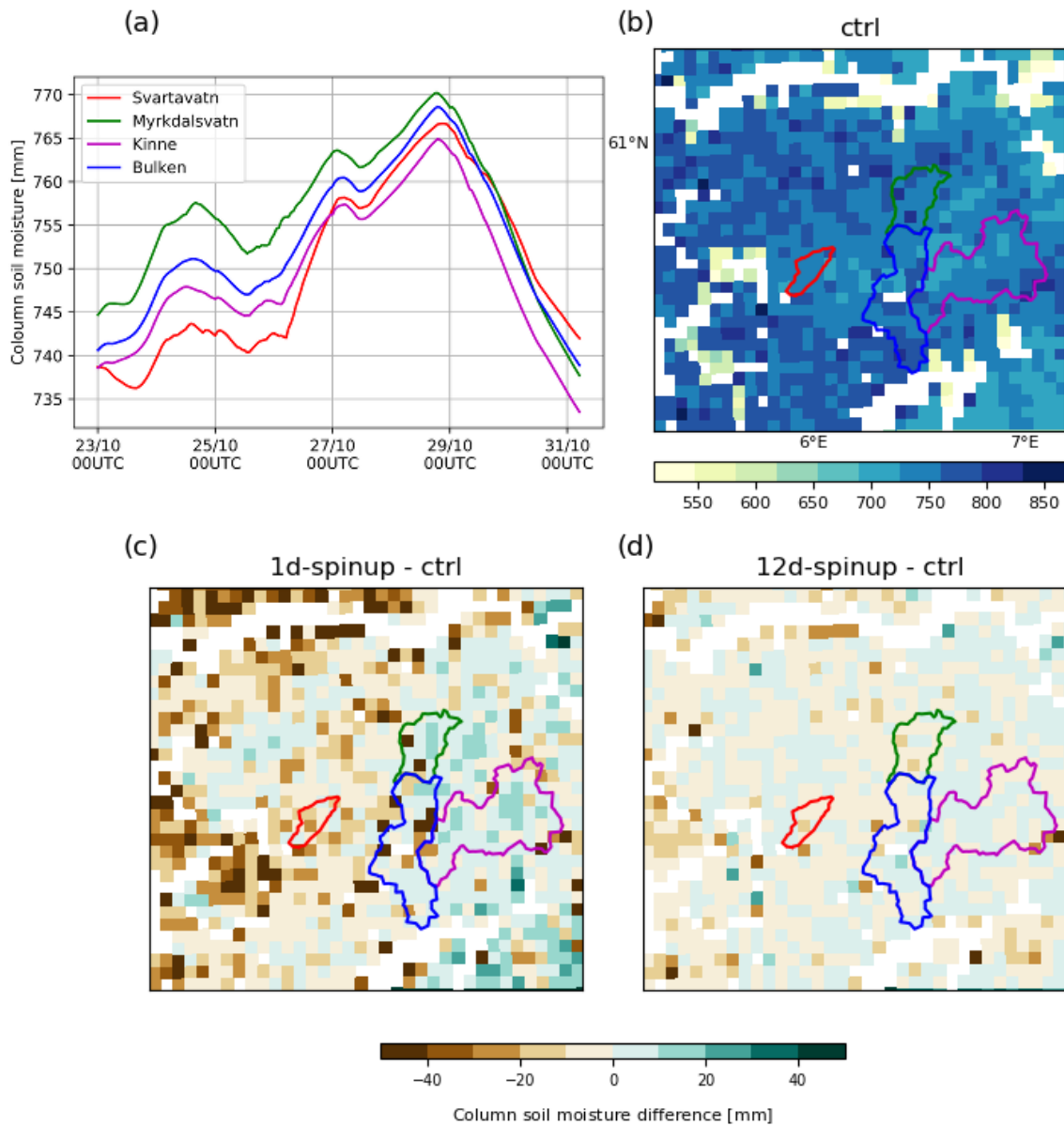



Figure 7: a) Temporal evolution of total column soil moisture averaged in the catchments (ctrl simulation); b) Averaged total column soil moisture on 24 October; c) soil moisture differences between the 1-day spinup and the control simulation; d) soil moisture difference between the 12-day spinup and the control simulation.

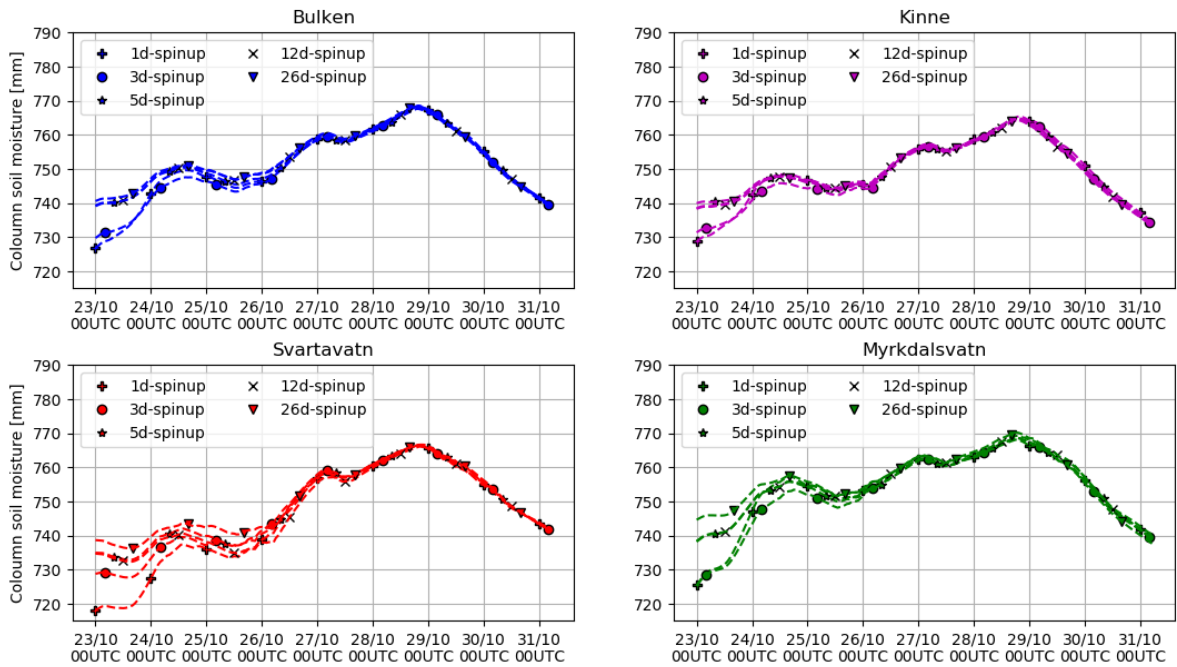



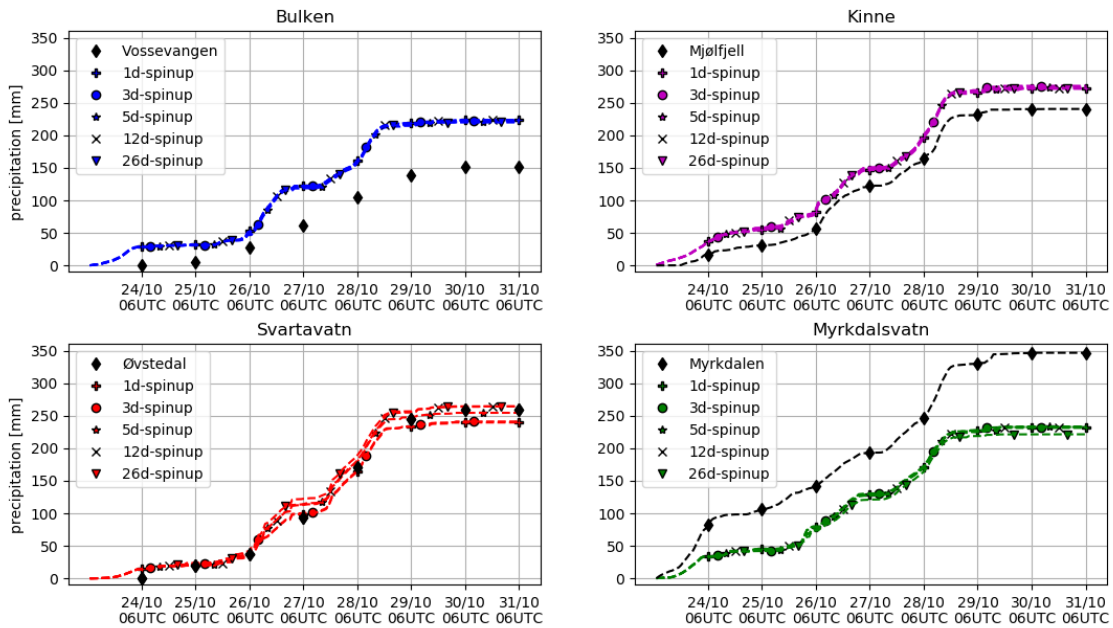
Figure 8: The four basin averaged simulated soil moisture with different spinup time, including 1 day (1d-spinup), 3 days (3d-spinup), 5 days (5d-spinup), 12 days (12d-spinup) and 26 days (26d-spinup) for the flooding events during 23 - 31 October, 2014.

750

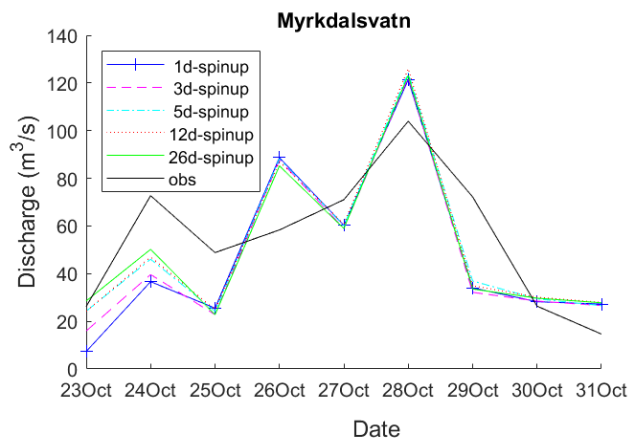
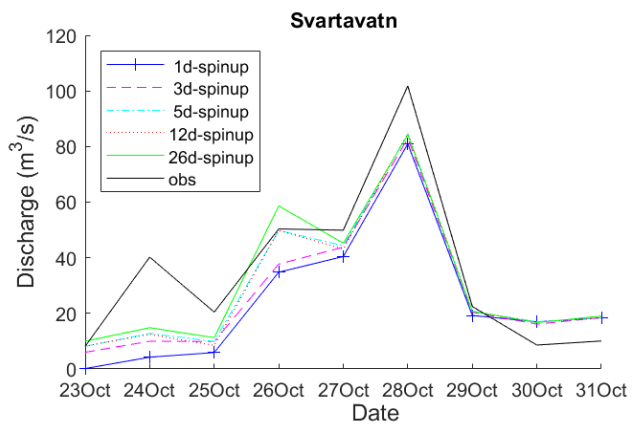
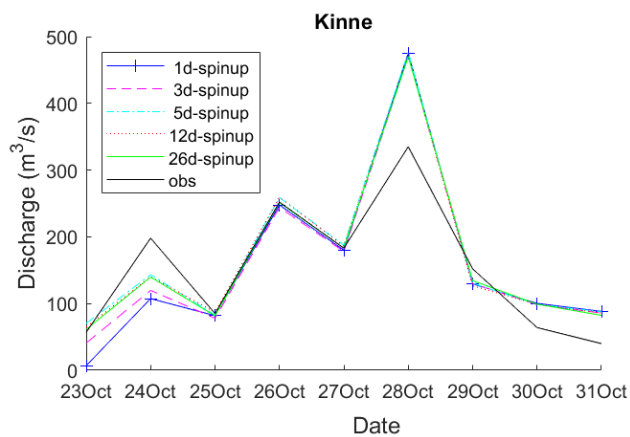
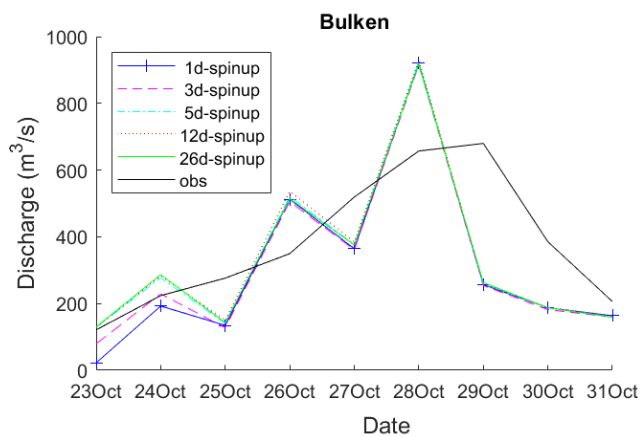
Figure 9: The accumulated precipitation from the simulations at the four nearest grid points to Vossevangen (Bulken), Mjølfjell (Kinne), Myrkdalen (Myrkdalsvatn) and Øvstedal (Svartavatn), observational stations from the simulations of different spinup times, including 1-day (1d-spinup), 3-day (3d-spinup), 5-day (5d-spinup), 12-day (12d-spinup) and 26-days (26d-spinup) compared with the observational station within the catchment.

755

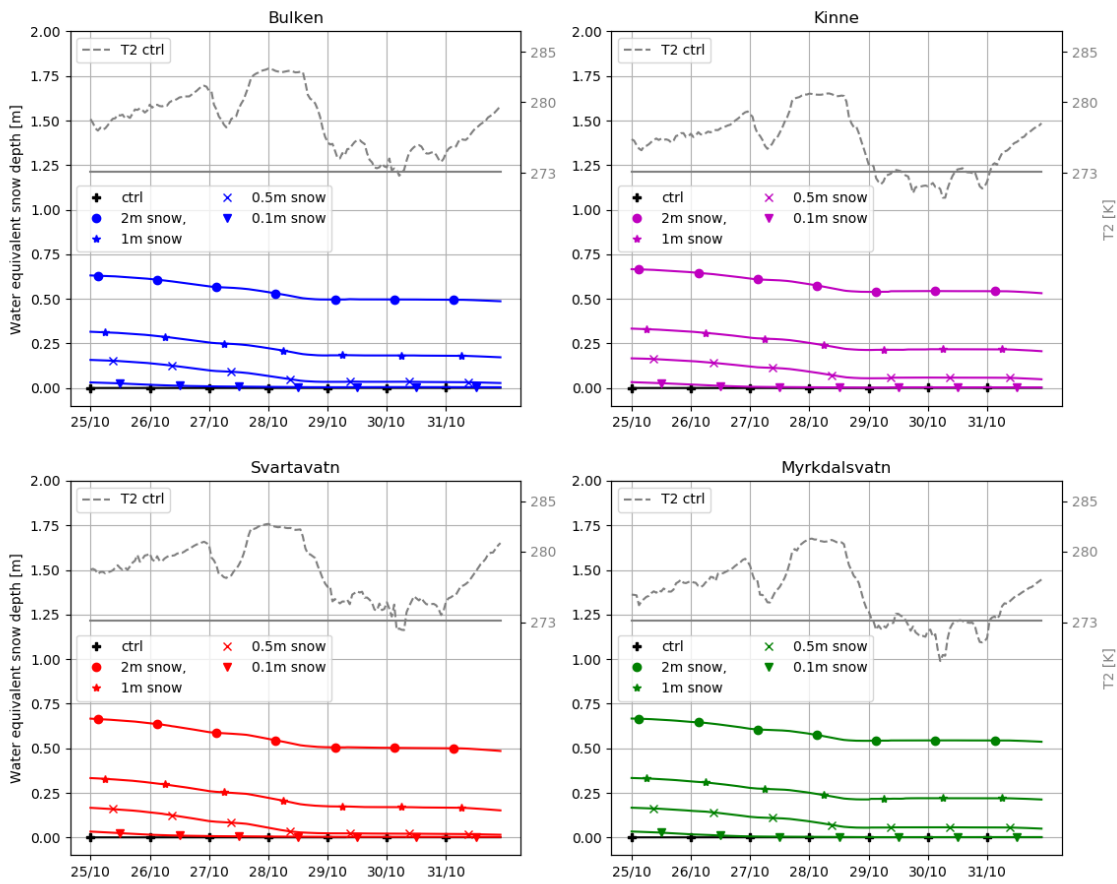
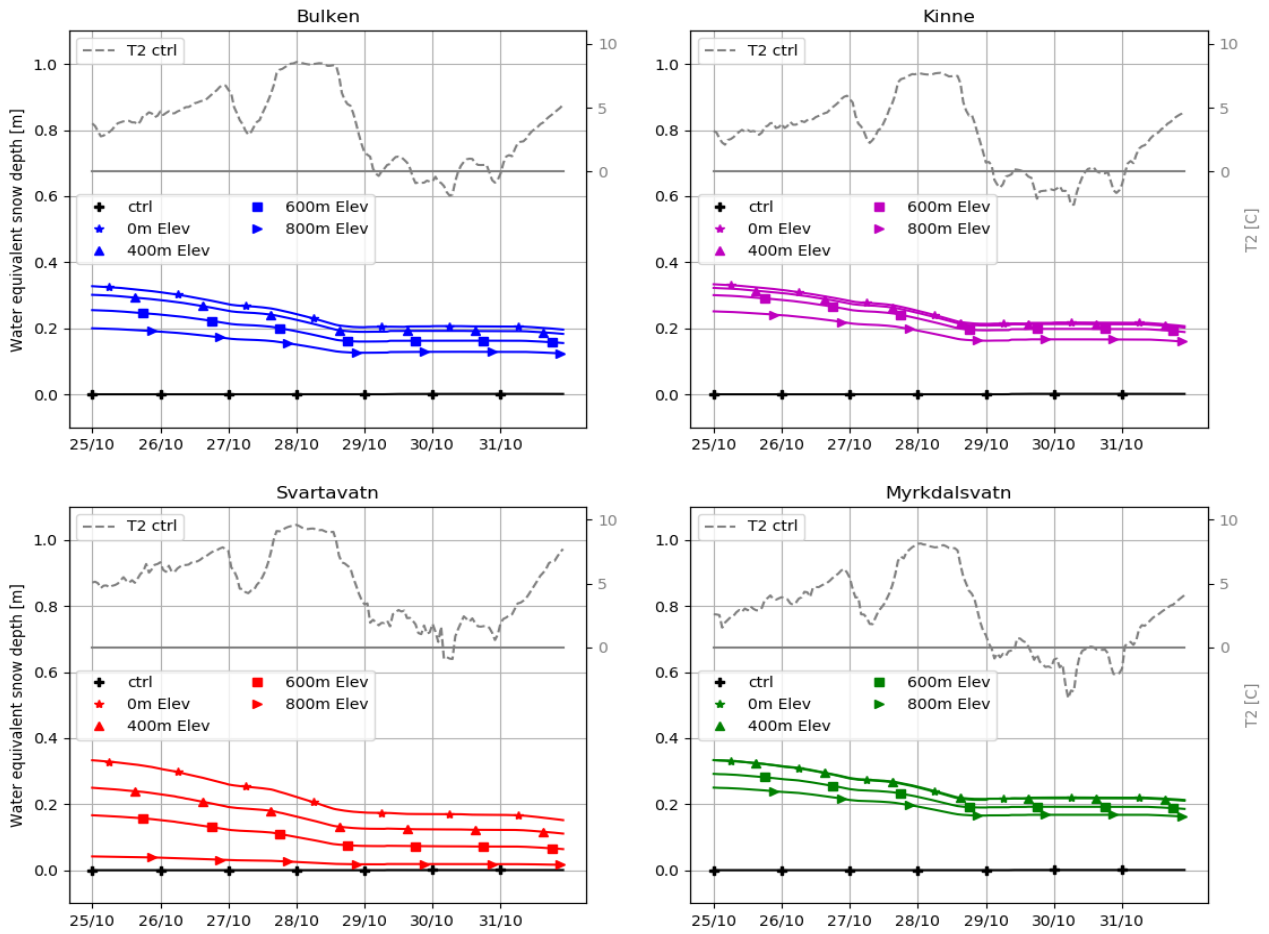
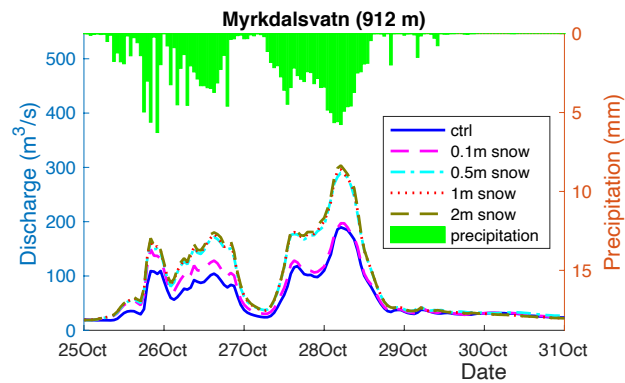
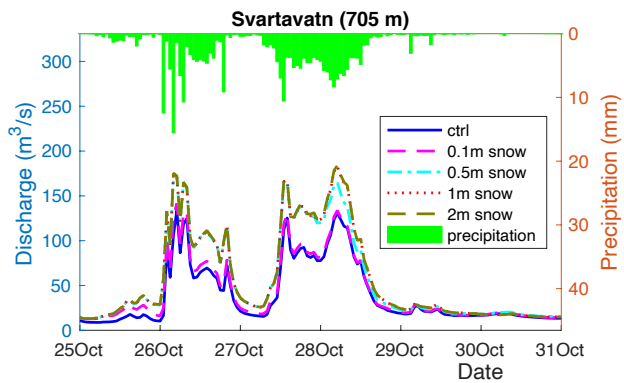
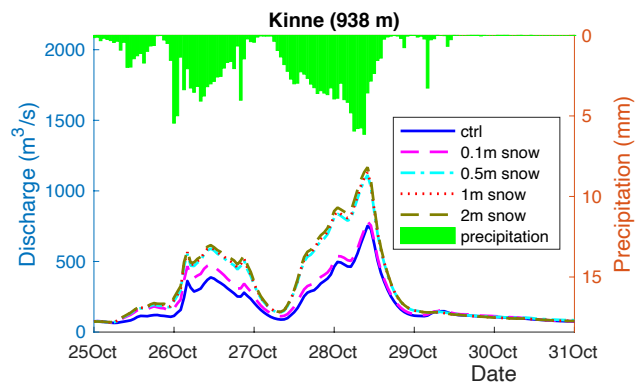
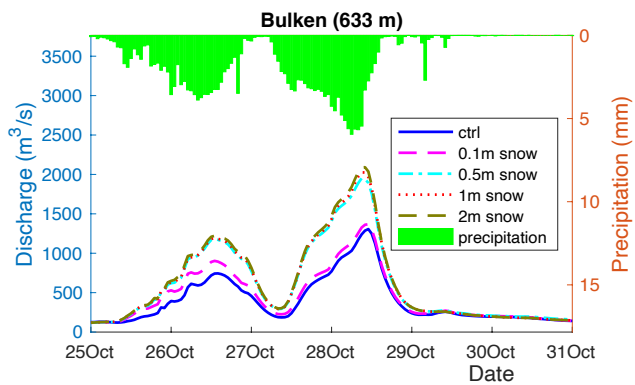
Figure 10: Daily observed streamflow (obs) and simulations discharges with different spinup time, including 1 day (1d-spinup), 3 days (3d-spinup), 5 days (5d-spinup), 12 days (12d-spinup) and 26 days (26d-spinup) for the flooding events during 23 - 31 October, 2014.

760 **Figure 11: Catchment averaged hourly snow water equivalents from 3km WRF-Hydro simulation for the snow sensitivity experiments, including 2 m added snow depth (2m snow), 1 m added snow depth (1m snow), 0.5 m added snow depth (0.5m snow), 0.1 m added snow depth (0.1m snow), and the catchment averaged hourly snow water equivalent (ctrl) and surface temperature (T2) from the control simulation of the 26 day spinup.**

765 **Figure 12: Catchment averaged snow water equivalents from 3km WRF-Hydro simulation for the pre-existing snow cover sensitivity experiments, including 1 m added snow depth to the area where the elevation is above 0 m (0m Elev), 400 m (400m Elev), 600 m (600m Elev), and 800 m (800m Elev) based on the CTRL simulation (26 day spinup)), and the catchment averaged hourly snow water equivalent (ctrl) and surface temperature (T2) from the control simulation of the 26 day spinup.**

770

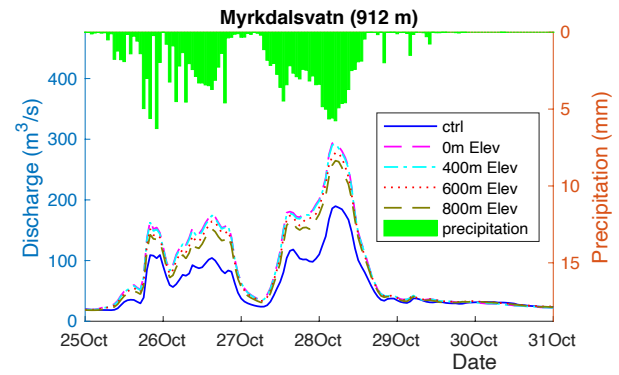
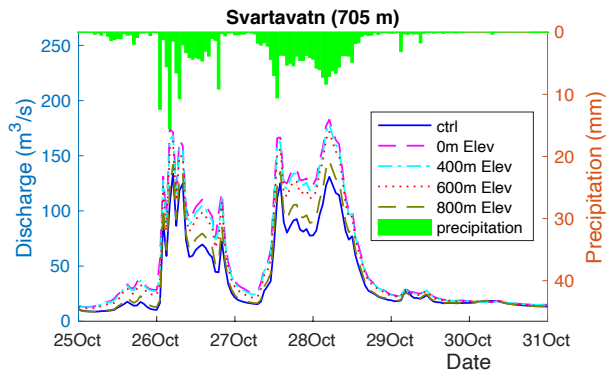
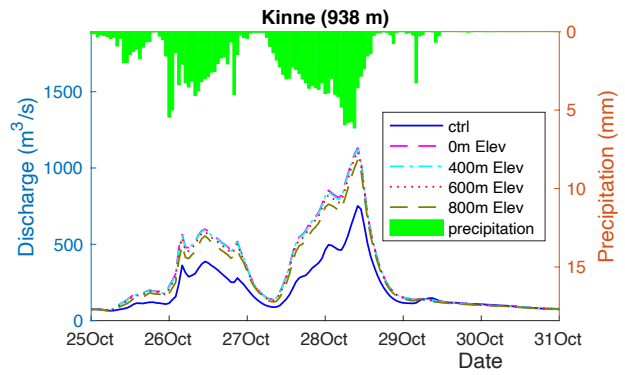
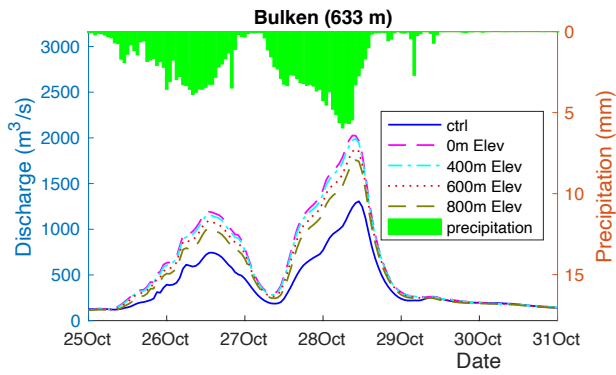
Figure 13: Hourly simulated discharge during 25 - 31 October from different experiments of added snow depths, i.e. 0.5 m, 1 m and 2 m snow and the control simulation from 26-day spinup without added snow depth (ctrl) (mean elevations of four basins are given in the subtitles).

775

Figure 14: Hourly simulated discharge during 25 - 31 October from the control simulation (ctrl) and the different experiments of added 1m snow depths on the areas with elevation above 0, 400 m, 600 m and 800 m (mean elevations of four basins are given in the subtitles).