HAL open science

Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts

Louis Quéno, Vincent Vionnet, Ingrid Dombrowski-Etchevers, Matthieu Lafaysse, Marie Dumont, Fatima Karbou

To cite this version:

Louis Quéno, Vincent Vionnet, Ingrid Dombrowski-Etchevers, Matthieu Lafaysse, Marie Dumont, et al.. Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts. The Cryosphere, 2016, 10 (4), pp.1571-1589. 10.5194/tc-10-1571-2016 . hal-03684126

HAL Id: hal-03684126

https://hal.science/hal-03684126

Submitted on 2 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

The Cryosphere, 10, 1571–1589, 2016
www.the-cryosphere.net/10/1571/2016/
doi:10.5194/tc-10-1571-2016
© Author(s) 2016. CC Attribution 3.0 License.

The Cryosphere
Open Access

Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts

Louis Quéno[1], Vincent Vionnet[1], Ingrid Dombrowski-Etchevers[2], Matthieu Lafaysse[1], Marie Dumont[1], and Fatima Karbou[1]

[1]Météo France/CNRS, CNRM UMR3589, CEN, St. Martin d'Hères, France
[2]Météo France/CNRS, CNRM UMR3589, Toulouse, France

Correspondence to: Louis Quéno (louis.queno@meteo.fr)

Received: 22 January 2016 – Published in The Cryosphere Discuss.: 4 February 2016
Revised: 23 June 2016 – Accepted: 4 July 2016 – Published: 22 July 2016

Abstract. Distributed snowpack simulations in the French and Spanish Pyrenees are carried out using the detailed snowpack model Crocus driven by the numerical weather prediction system AROME at 2.5 km grid spacing, during four consecutive winters from 2010 to 2014. The aim of this study is to assess the benefits of a kilometric-resolution atmospheric forcing to a snowpack model for describing the spatial variability of the seasonal snow cover over a mountain range. The evaluation is performed by comparisons to ground-based measurements of the snow depth, the snow water equivalent and precipitations, to satellite snow cover images and to snowpack simulations driven by the SAFRAN analysis system. Snow depths simulated by AROME–Crocus exhibit an overall positive bias, particularly marked over the first summits near the Atlantic Ocean. The simulation of mesoscale orographic effects by AROME gives a realistic regional snowpack variability, unlike SAFRAN–Crocus. The categorical study of daily snow depth variations gives a differentiated perspective of accumulation and ablation processes. Both models underestimate strong snow accumulations and strong snow depth decreases, which is mainly due to the non-simulated wind-induced erosion, the underestimation of strong melting and an insufficient settling after snowfalls. The problematic assimilation of precipitation gauge measurements is also emphasized, which raises the issue of a need for a dedicated analysis to complement the benefits of AROME kilometric resolution and dynamical behaviour in mountainous terrain.

1 Introduction

A major challenge in seasonal snow cover studies in mountainous terrain is to take into account the high spatial variability of the snowpack, since it affects many phenomena in mountains. In particular, it is of prime importance for avalanche hazard forecasting or mountain hydrology. The snow cover heterogeneous distribution is indeed the main factor controlling the runoff during the melting season (Anderton et al., 2002), as well as an essential factor of avalanche formation (Schweizer et al., 2003). The seasonal snow heterogeneity also strongly affects the alpine tundra plant life (Jonas et al., 2008b), as well as the alpine wildlife (Jonas et al., 2008a).

The spatial variability of the snowpack is observed at different scales and is mainly caused by the spatial variability of atmospheric conditions, on the same range of scales. The regional climate determines the main synoptic weather patterns which contribute to the snow cover buildup. Within a mountain range and at a given elevation, the snowpack spatial variability is caused by the amount of local exposure to synoptic flows bringing snowfall. Additionally, the atmospheric conditions at the surface vary following the local topography, e.g. the elevation influences temperatures, precipitation phase and radiations, and slope and aspect have an influence on incoming solar radiations. At a smaller scale (less than 100 m), processes like wind-induced erosion (Pomeroy and Gray, 1995), avalanches (Schweizer et al., 2003) or preferential deposition of snowfall on the leeward slopes (Lehning et al., 2008) play a decisive role on the snow distribution (e.g. Mott et al., 2010).

The description of the snowpack variability through snowpack modelling is thus highly dependent on the spatial resolution of the atmospheric forcing. This variability is currently represented by classes of elevation, slope and aspect at a scale of about 1000 km^2 for operational avalanche hazard forecasting in French mountainous areas. The detailed snowpack model SURFEX/ISBA/Crocus (Vionnet et al., 2012), mentioned as Crocus hereafter, is used within the SAFRAN–SURFEX/ISBA/Crocus–MEPRA model chain (Durand et al., 1999; Lafaysse et al., 2013). The meteorological analysis and forecasting system SAFRAN (Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige (Analysis System Providing Atmospheric Information to Snow); Durand et al., 1993) provides relevant meteorological parameters affecting the snowpack evolution, with a dependence on the elevation within mountain ranges, so-called "massifs", assumed to be homogeneous from a meteorological viewpoint. SAFRAN was also used in many other applications such as a climatology of the snow cover in the French Alps from 1958 to 2005 (Durand et al., 2009a, b).

The atmospheric forcing of snowpack models for distributed simulations (i.e. on a regular grid) has been recently the object of many studies, building on the development of NWP (numerical weather prediction) models of increasing resolution. Bellaire et al. (2011, 2013) performed snowpack simulations in Canada with the detailed snow cover model SNOWPACK (Bartelt and Lehning, 2002), driven by the 15 km resolution regional NWP model GEM15 (Mailhot et al., 2006), with a view to forecasting avalanche hazard. They highlighted that distributed snow cover simulations driven by NWP systems would be highly beneficial in areas with few snow cover observations. For snowpack simulations in mountainous terrain, kilometric atmospheric information allows us to capture an important part of the intra-massif snowpack variability. Such simulations were performed by Bellaire et al. (2014) in New Zealand for avalanche hazard forecasting, driving SNOWPACK by the NWP model ARPS (Advanced Regional Prediction System; Xue et al., 2000) at a 3 and 1 km horizontal resolution. This study shows better results in terms of snowfall for the highest resolution forcing over a 10-day snowy period. Horton et al. (2015) demonstrated the benefits of forcing SNOWPACK with the 2.5 km resolution NWP model GEM-LAM (Erfani et al., 2005) for specific studies of snowpack stability (surface hoar layers formation). Schirmer and Jamieson (2015) applied the same chain of models GEM-LAM/SNOWPACK in the mountains of Western Canada and north-western USA, with a focus on winter precipitation, and showed that the kilometric-resolution NWP system performed better than GEM15 (15 km) and a precipitation analysis system, particularly in terms of snowfall quantitative distribution. The snowpack variability can also be simulated at scales of tens of metres, using adequate snowpack–atmosphere coupled models. Vionnet et al. (2014) used the coupled system Meso-NH/Crocus to study wind-induced erosion of the snowpack, at a 50 m horizontal resolution, and Mott et al. (2014) used the atmospheric model ARPS at a 75 m horizontal resolution to study the orographic effects on snow deposition patterns. Such simulations can only be made on very limited areas, due to obvious computing limitations, and cannot currently be applied to operational issues such as avalanche hazard forecasting or mountain hydrology.

The aim of the present study is to simulate the snowpack variability within a whole mountainous chain. Consequently, kilometric snowpack simulations offer a promising compromise between spatial resolution and computational time. AROME (Application of Research to Operations at MEsoscale; Seity et al., 2011) is a 2.5 km resolution NWP model, operational over France since December 2008. Its kilometric resolution over the French mountains offers an alternative to the forcing of Crocus by SAFRAN, at higher resolution, but without a dedicated analysis system. AROME has been preliminarily evaluated in mountainous terrain by Dombrowski-Etchevers et al. (2013) and Vionnet et al. (2016), who showed its good performance for mountain weather forecast in the French Alps. Vionnet et al. (2016) discussed the potential of AROME–Crocus for snowpack modelling in the French Alps. They illustrated the realistic representation of the intra-massif spatial variability of the snowpack for this region, although the improved resolution does not compensate for the lack of a dedicated analysis system. Subsequently, this paper proposes to expand the study to the French and Spanish Pyrenees, whose climate differs from that of the Alps as these mountains are subjected to the influence of both the Atlantic Ocean and Mediterranean Sea. We also refine the analysis of snowpack simulations, using categorical scores to separate the different physical processes.

The organization of the paper is as follows. In Sect. 2, we introduce briefly the geographical and climate characteristics of the study area and period. Section 3 describes the snowpack model Crocus, then the atmospheric forcing from NWP model AROME at kilometric resolution and the forcing from SAFRAN reanalysis, and, finally, the observations dataset and verification methods. Section 4 details the results following three main axes: (i) global scores and spatial distribution of snow depth (SD), (ii) daily snow depth variations and winter precipitation and (iii) comparison to snow water equivalent (SWE) scores and study of bulk snowpack density. These results are discussed in Sect. 5, with concluding remarks and outlooks.

2 Study area and period

This study focusses on the Pyrenees (Fig. 1), the natural border which separates France from Spain, from the Atlantic Ocean to the Mediterranean Sea. Many summits, especially in its central part, exceed 3000 m a.s.l. with a maximum at the Aneto Peak in Spain with 3404 m a.s.l. Our domain of study

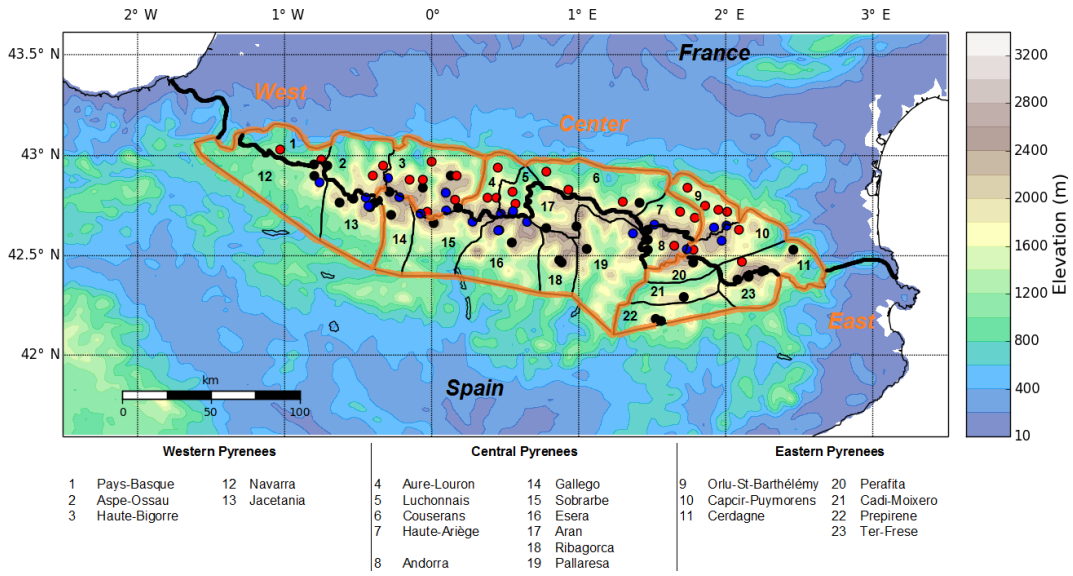

Figure 1. Location of measurement stations in the Pyrenees: SD and precipitation (red circles), SD and SWE (blue circles) and SD only (black circles). Background map: AROME topography (years 2010–2012). SAFRAN massifs delimited (black line), national borders (bold black line) and climatic regions (bold orange line). SAFRAN massifs names in caption.

covers France, Andorra and Spain, from 41.6 to 43.6° N latitude and from −2.5 to 3.5° E longitude (approximately 500 km × 220 km).

The Pyrenean climate, in its western part, is strongly influenced by the proximity of the Atlantic Ocean and therefore mostly exposed to westerly winds. This influence abates in the eastern Pyrenees. Hence, most winter precipitations controlling the snow cover distribution are due to south-western to north-western flows (e.g. Buisan et al., 2015; Durand et al., 2012; Maris et al., 2009; Vada et al., 2013). They generate a strong west–east gradient of decreasing precipitation, leading to a similar gradient of mean snow depth and of number of days with snow on the ground (Maris et al., 2009). A north–south gradient of snow quantities (with more snow on the northern side) is due to warmer and drier conditions in Spain than in France, largely associated with a frequent northerly Foehn effect in Spain (López-Moreno et al., 2009). Following Maris et al. (2009), we defined three climatic regions: western Pyrenees, under the direct influence of the Atlantic Ocean, central Pyrenees, with a more continental climate, and eastern Pyrenees, under the Mediterranean influence (Fig. 1).

The study period goes from August 2010 to July 2014. Because of the interannual variability of winter conditions, several years are necessary to assess snow models with significance (Essery et al., 2013). Moreover, the 2010–2014 period covers four very contrasted winters. Winter 2010–2011 was rather dry, hence a deficit of snow in the Pyrenees (with respect to the climate normal), despite early snowfall in November. Winter 2011–2012, also dry, saw a deficit of snow, especially on the Spanish slopes (Vada et al., 2013; Gascoin et al., 2015). In contrast, winter 2012–2013 was very cold and wet, breaking a 40-year old record of snowfall and snow depth, particularly in the French Pyrenees. Winter 2013–2014 was also characterized by a much higher level of snow than normal due to a lot of precipitation, despite warmer conditions.

3 Data and methods

3.1 Snowpack model

Snowpack simulations were carried out using the detailed snow cover model Crocus (Brun et al., 1992; Vionnet et al., 2012) coupled with the ISBA land surface model within the SURFEX (EXternalized SURFace) simulation platform (Masson et al., 2013). SURFEX/ISBA/Crocus models the evolution of the physical properties of the snowpack, its stratigraphy (with a user-defined maximum number of layers – 50 in this study) and the underlying ground, under given meteorological forcing data. The model is used here in an offline mode (i.e. not fully coupled to atmospheric simulations), with prescribed atmospheric forcing described in Sect. 3.2. Snowpack simulations were performed over the domain defined in Sect. 2 (Fig. 1), on a regular 0.025° grid, from 1 August 2010 to 31 July 2014, with a 15 min internal time step.

Soil properties were obtained from the HSWD (Harmonized World Soil Database) 1 km resolution database for soil texture (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012). Aspect and slope are not taken into account for incoming solar radiations, since the 2.5 km resolution topography can hardly

represent the local orography of observation stations. As observations are collected in open fields, the interactions with the vegetation and the parameterization of fractional snow cover are not activated within the SURFEX scheme. Wind-induced snow transport is not simulated.

3.2 Atmospheric forcing

Crocus requires the following atmospheric forcings: reference level temperature and specific humidity (usually 2 m above ground), wind speed (usually 10 m above ground), incoming short-wave and long-wave radiations, and solid and liquid precipitation. Two different forcings were used: one generated from the AROME NWP system (Seity et al., 2011) operational forecasts and the other one from the SAFRAN reanalyses (Durand et al., 1993, 2009b). These forcings are described hereafter.

3.2.1 AROME: kilometric-resolution NWP system

AROME is the high-resolution NWP system at Météo France (Seity et al., 2011). Its 2.5 km horizontal resolution (upgraded to 1.3 km in 2015; Brousseau et al., 2016) makes it of particular interest for forecasting intense events (like convective rains) and small-scale processes in alpine terrain, such as orographic precipitations or Foehn effects, thanks to a realistic description of the topography. AROME is a spectral and non-hydrostatic model which combines the physical package of the research model Meso-NH (Lafore et al., 1998) with the dynamical core of the non-hydrostatic version of the limited area NWP ALADIN model (Bubnová et al., 1995). A detailed description of the physics and data assimilation schemes can be found in Seity et al. (2011). In particular, the precipitation phase is derived from the cloud microphysical scheme.

The implementation of AROME as an operational system is made through 30 h forecasts at the 00:00, 06:00, 12:00 and 18:00 UTC nominal analysis times, over a domain covering France. We use here the hourly forecasts issued from the 00:00 UTC analysis time, from +6 to +29 h, extracted on a regular latitude / longitude 0.025° grid to build a continuous forcing from 1 August 2010 to 31 July 2014 over the domain of study.

Some changes in the operational configuration of AROME occurred during the 4 years of simulations: the simulation domain was extended during summer 2012 with a modification of the topographic database. The topography from Global 30 Arc-Second Elevation dataset (GTOPO30) was used in a low-resolution version (5 km) before summer 2012 and at 30 arcsec (approximately 1 km) resolution afterwards, which led to a modification of the forcing files orography in the middle of our simulation period.

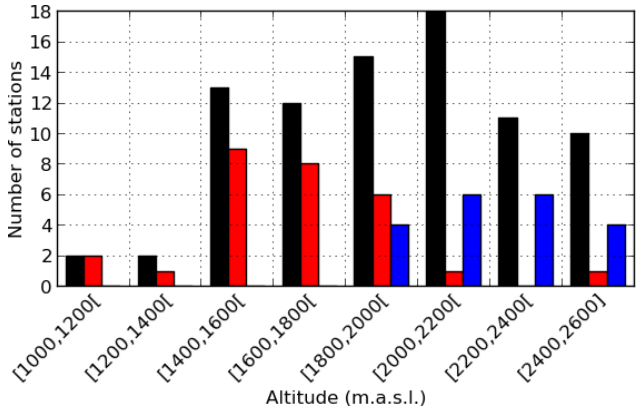

Figure 2. Altitude distribution of all SD stations (black), precipitation gauges (red) and SWE stations (blue).

3.2.2 SAFRAN: analysis system

The SAFRAN analysis system (Durand et al., 1993, 2009a, b) provides hourly atmospheric forcing data for each of the 23 massifs of the Pyrenees (Fig. 1). Within each massif, the forcing is provided by 300 m altitude steps. SAFRAN reanalyses take a preliminary guess from the global NWP model ARPEGE (from Météo France, 15 km grid spacing guess projected on a 40 km grid), complemented by available observations from automatic weather stations, manual observations carried out in the climatological network and in ski resorts and atmospheric upper-level sounding. In particular, a daily precipitation analysis is included, with a climatological guess depending on a daily determination of the general weather pattern. This determination is based on a classification of nine weather patterns, defined by Météo France mountain forecasters to be representative of the main precipitating regimes of the Pyrenees. It is made following the synoptic circulation, through the altitude of the 500 hPa geopotential level. The precipitation phase is derived from a simple threshold of 1 °C air temperature at 2 m above the ground. In this study, SAFRAN forcing was interpolated over the 0.025° grid of the domain described in Sect. 2, following the method described by Vionnet et al. (2012).

3.3 Evaluation dataset

The observational dataset contains SD, SWE and precipitation measurements available in the Pyrenean SAFRAN massifs, both in France and Spain. The SD observations consist of daily manual measurements at ski resorts (at 06:00 UTC) and hourly automatic measurements by ultra-sonic sensors at high-altitude stations. Only the value at 06:00 UTC from the hourly record is used in this study. The SWE measurements come from automatic stations with cosmic ray snow gauges (Gottardi et al., 2013). Daily values are obtained through a 24 h median smoothing of hourly measurements. Both SD and SWE data are independent (i.e. not assimi-

Table 1. Two-by-two contingency table.

	OY	ON
FY	HI (hits)	FA (false alarms)
FN	MI (misses)	CR (correct rejections)

OY is observed yes; ON is observed no; FY is forecast yes; FN is forecast no.

lated in SAFRAN–Crocus nor in AROME–Crocus). The 24 h cumulated precipitations measurements are manually collected every day at ski resorts with precipitation gauges (at 06:00 UTC), without any correction. These data are assimilated in SAFRAN.

A criterion of altitude is then applied to select adequate stations. Only stations with less than 150 m elevation difference to the model topography are selected for evaluation. Following this selection, 83 SD stations could be used in the whole Pyrenees, amongst which 20 stations with SWE measurements and 28 stations with precipitation measurements (Fig. 2); 45 of them are located in France, 38 in Spain, 24 in the western Pyrenees, 32 in the central Pyrenees and 17 in the eastern Pyrenees (Fig. 1). These stations are all between 1000 and 2600 m a.s.l. The altitude distribution is represented in Fig. 2. The mean altitude, weighted by the number of SD observations, is 2007 m a.s.l. The spatial coverage of the domain can be considered representative (observations are available for all massifs), excepting the southern foothills with no data.

MODIS daily fractional snow cover images (MOD10A1, Klein and Stroeve, 2002) at 0.005° resolution are used to evaluate the ability of snowpack simulations to reproduce the spatial variability of snow cover in the Pyrenees. They are projected to a 0.025° grid using a nearest-neighbour interpolation method for systematic comparison to snow cover simulations.

3.4 Evaluation methods

AROME–Crocus snowpack simulations were evaluated in terms of SD and SWE from 1 October to 30 June over the period 2010–2014. SAFRAN–Crocus simulations were evaluated in a similar manner. Two error metrics were used: the bias and the standard deviation error (STDE, which represents the temporal and spatial dispersion around the bias).

A complementary evaluation was carried out in terms of daily snow depth variations. This additional metrics allows us to avoid cumulative errors which occur during winter and to offer another view on precipitation forecast as well as the simulation of settling and ablation processes. The daily snow depth variation $\Delta\mathrm{SD}_n$ is defined for day n as

$$\Delta\mathrm{SD}_n = \mathrm{SD}_n - \mathrm{SD}_{n-1}. \tag{1}$$

ΔSD categories are defined according to the decrease or increase of SD and allow to study categorical distribution, sums and scores, in a similar way as Schirmer and Jamieson (2015) in their study of winter precipitations. Daily snow water equivalent variation (ΔSWE) is also defined in the same way.

Based on two-by-two contingency tables (Table 1), the Equitable Threat Score (ETS; defined by Nurmi, 2003) was used to study daily variations. The ETS is a score commonly used for precipitation forecast evaluation (e.g. Bélair et al., 2009). It was used here for the purpose of comparison with the findings of Schirmer and Jamieson (2015). It measures the proportion of correct "yes" events amongst all events, except correct rejections (the forecast skill does not consider "no" events, which are much more frequent than "yes" events):

$$\mathrm{ETS} = \frac{\mathrm{HI} - \mathrm{HI}_{\mathrm{rdm}}}{\mathrm{HI} + \mathrm{FA} + \mathrm{MI} - \mathrm{HI}_{\mathrm{rdm}}}. \tag{2}$$

Taking into account chance hits,

$$\mathrm{HI}_{\mathrm{rdm}} = \frac{(\mathrm{HI} + \mathrm{FA})(\mathrm{HI} + \mathrm{MI})}{N}, \tag{3}$$

where $N = \mathrm{HI} + \mathrm{FA} + \mathrm{MI} + \mathrm{CR}$ is the total number of observations. It ranges from $-1/3$ to 1, where 0 means no skill and 1 means perfect score.

The Jaccard index (J) and the average symmetric surface distance (ASSD) are two similarity metrics which were used to compare simulated and remotely sensed snow-covered areas. They were calculated with the Python medpy.metric.binary program from the MedPy package. They were applied to simulated and observed binary snow-covered maps on the same grid. If A and B represent the simulated and the observed snow cover domain, respectively, J is the number of pixels that are snow covered in both A and B divided by the total number of pixels in the union of A and B:

$$J = \frac{|A \cap B|}{|A \cup B|}. \tag{4}$$

J is thus dependent on the whole snow-covered area. It ranges from 0 to 1, where 0 means no overlap of A and B surfaces, and 1 means $A = B$. The ASSD is complementary to J since it evaluates a mean distance between the boundaries of the two surfaces. It is based on the modified directed Hausdorff distance between boundaries L_A and L_B, defined by Dubuisson and Jain (1994) as the average distance of the points of L_A to L_B:

$$\mathrm{MDHD}(A, B) = \frac{1}{|L_A|} \sum_{a \in L_A} \mathrm{d}(a, L_B), \tag{5}$$

where $\mathrm{d}(a, L_B)$ is the Euclidean distance between point a and the closest point of boundary L_B:

$$\mathrm{d}(a, L_B) = \inf_{b \in L_B} ||a - b||. \tag{6}$$

The MDHD is a directed distance, used by Sirguey (2009) for snow pattern matching. The ASSD is its symmetrised version:

$$\mathrm{ASSD}(A, B) = \frac{\mathrm{MDHD}(A, B) + \mathrm{MDHD}(B, A)}{2}. \quad (7)$$

It ranges from 0 to $+\infty$, where 0 means $L_A = L_B$. In practice, the maximum value is the highest possible distance between two points of the domain.

Binary maps are built using a 20 mm SWE threshold for simulations and a 50 % snow fraction threshold for satellite data. The metrics are calculated only when the cloud fraction on the domain is less than 10 % and the snow cover represents at least 10 pixels in MODIS images interpolated on AROME grid (the size of a pixel is $0.025° \times 0.025°$, i.e. approximately 6.25 km^2).

4 Results

4.1 Evaluation of simulated snow depth

4.1.1 Global scores for the winter season

Table 2 summarises error statistics for snow depth during the whole period of study. The number of stations available varies from year to year (from 62 to 79) because of modifications in the model topography and missing data. Scores were also computed for a constant number of stations (restricted to 46, not shown) and showed that the annual variability of the number of stations does not impact the results and the analysis exposed hereafter. These scores show a global overestimation of snow depth by AROME–Crocus with an overall bias of +55 cm, while the overall bias of SAFRAN–Crocus is +22 cm. The overall STDE reaches 70 cm for AROME–Crocus compared to 57 cm for SAFRAN–Crocus. The errors are rather high for both models, which will be explained in the next sections.

For both models, the highest STDEs are found for winters 2012–2013 and 2013–2014, two very snowy winters. In terms of spatial distribution, the positive bias and STDE decrease from west to east for AROME–Crocus, with notable errors in the western zone. In the eastern zone, AROME–Crocus and SAFRAN–Crocus STDEs are equivalent. AROME–Crocus scores are equivalent in France and Spain, while SAFRAN–Crocus behaves slightly better in France, probably due to a higher number of observations assimilated by the model. In regard to altitude, biases are constant for SAFRAN–Crocus and decrease for AROME–Crocus, which implies a higher relative bias in the [1000 m, 1800 m[range.

Figure 3 shows scores for each station over the whole period of study. Almost all stations show an overestimation of snow depth, particularly for AROME–Crocus with extreme positive biases on the Atlantic foothills. The three highest biases for AROME–Crocus are given by the following three stations: Isaba El Ferial (+188 cm; massif of Navarra, western Pyrenees, Spain), Arette La Pierre Saint Martin (+209 cm; massif of Pays-Basque, western Pyrenees, France) and Soum Couy Nivôse (+229 cm; massif of Aspe–Ossau, western Pyrenees, France), all located in the vicinity of the Pic d'Anie, the first summit above 2500 m a.s.l. close to the Atlantic Ocean. These three stations also show a very high STDE (higher than 1 m). The two next highest biases are located in the north-west foothills: Gourette (+135 cm; massif of Aspe–Ossau, western Pyrenees, France) and Hautacam (+154 cm; massif of Haute Bigorre, western Pyrenees, France). This region is particularly exposed to W-NW flows due to its proximity to the Atlantic Ocean. There is thus an excessive orographic blocking on these first peaks by AROME. Except for these stations, biases and STDEs are more homogeneous in the rest of the Pyrenees.

4.1.2 Focus on winter 2011–2012

Winter 2011–2012 was characterized by a deficient snowpack in the Spanish Pyrenees, due to dry and warm weather in the southern side of the chain (Vada et al., 2013). It was also characterized by a strong contrast between the French and the Spanish sides of the Pyrenees: even if the French Pyrenees exhibited a deficit of snow for most of the winter (with respect to the climate normal), the first half of February 2012 was exceptionally cold and snowy in France. The Spanish Pyrenees were far less prone to snowfalls due to the northern flow. This asymmetry (and the ensuing fall in the Spanish hydropower production in springtime) was highlighted in terms of snow cover duration in Gascoin et al. (2015). Hereafter are shown the added value of AROME high-resolution forcing for simulating a particular meteorological contrast due to the topography and the resulting snow cover distribution.

Figure 4 gives an overview of the snow cover simulated by AROME–Crocus and SAFRAN–Crocus (values of SWE higher than 20 mm), compared to MODIS fractional snow cover images, on 22 February 2012. This date (selected because of clear sky conditions) is close to the end of the intense cold and snowy events in the French Pyrenees, corresponding to a maximum contrast between both sides of the Pyrenees. This contrast appears clearly on MODIS snow cover image, where snow is only present on the highest summits of the Spanish Pyrenees, on the border ridge, while snow covers most of the French Pyrenean massifs and Val d'Aran (in Spain, but on the northern side of the Pyrenean highest ridge). The absence of snow in the Spanish Pyrenean foothills is particularly well represented in the AROME–Crocus simulation, and the snow cover distribution matches observations. On the contrary, SAFRAN–Crocus simulation exhibits a rather homogeneous snow cover in Spanish massifs (despite still lower quantities than in the French Pyrenees). The snow cover spatial distribution, and particularly

Table 2. Scores (bias and STDE) for simulated snow depth against observations in the Pyrenees for winters 2010–2011 to 2013–2014.

	Stations	*N*	Mean obs. (cm)	Bias (cm) AROME	Bias (cm) SAFRAN	STDE (cm) AROME	STDE (cm) SAFRAN
2010–2014	83	47 169	70	55	22	70	57
2010–2011	63	10 445	48	57	20	55	42
2011–2012	62	10 401	39	43	16	52	44
2012–2013	79	14 281	103	52	17	77	65
2013–2014	67	12 042	76	65	37	85	64
West	27	14 393	83	65	17	84	54
Centre	35	21 865	72	57	28	64	55
East	21	10 911	50	36	18	58	63
France	45	22 491	76	56	17	75	50
Spain	38	24 678	65	53	28	66	62
[1000 m, 1800 m[	29	11 975	48	66	25	71	43
[1800 m, 2200 m[	33	19 164	76	46	17	72	61
[2200 m, 2600 m[	21	16 030	80	57	27	66	61

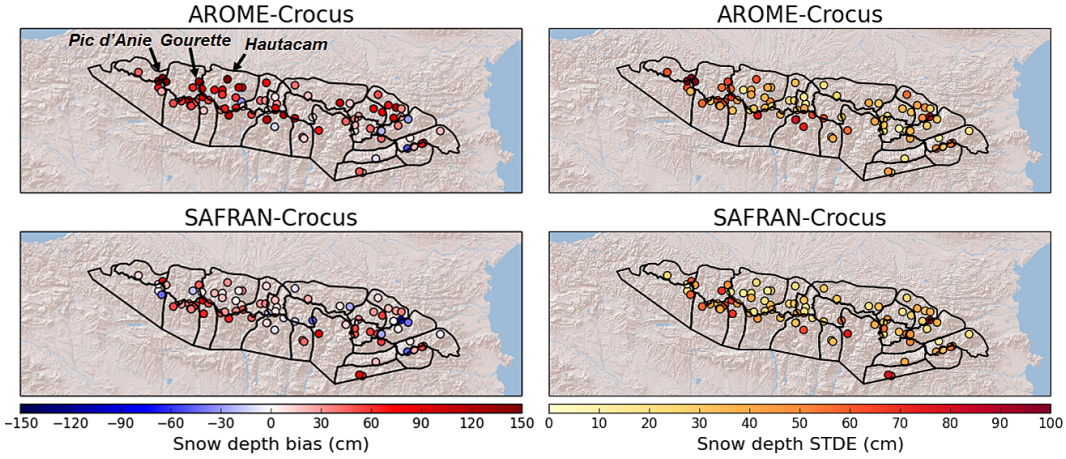

Figure 3. Snow depth bias (left) and STDE (right) by station for AROME–Crocus (up) and SAFRAN–Crocus (down), 2010–2014.

the snow deficit in the Spanish Pyrenees, is thus better simulated by AROME–Crocus.

This improvement in terms of snow cover may be attributed to AROME dynamical behaviour in complex topographies. Vada et al. (2013) showed that the snowfall deficit in 2011–2012 was more sensitive at Spanish stations exposed to southern flows, while Spanish stations more exposed to northern flows exhibited a lower negative anomaly. The snowpack was mainly constituted by N-NW flows during this season, which is confirmed by a study of SAFRAN weather patterns. We cumulated all snowfalls (from SAFRAN outputs) which occurred on the studied domain between 1 October 2011 and 22 February 2012 (date studied in Fig. 4). Of the cumulated snowfall, 67 % fell during days of north to north-western flows, which correspond to two synoptic patterns: a minimum geopotential in the Genoa gulf and a maximum in Ireland, associated with N and NW flows (38 %), and disturbed NW flow with strong geopotential gradient, implying strong precipitations on the NW French Pyrenees and a Foehn effect in Spain (29 %). During the four winters 2010–2014, these synoptic conditions constituted 45 % of total snowfalls. In contrast, only 4 % of total snow quantities fell during days of south to south-western flows (against 14 % over the period 2010–2014).

The behaviour of both forcing models in such specific synoptic conditions is of particular interest. Snowfalls from AROME and SAFRAN were cumulated from 1 October 2011 to 22 February 2012. They are represented in Fig. 5 along a NW/SE cross section, as well as cumulated positive ΔSWE from measurements of three stations close to the transect. Orographic blocking is visible on the windward sides, with a maximum snowfall immediately upstream of the highest summit whereas a Foehn effect in Spain implies a drastic drop of snowfalls immediately behind the highest ridge. The

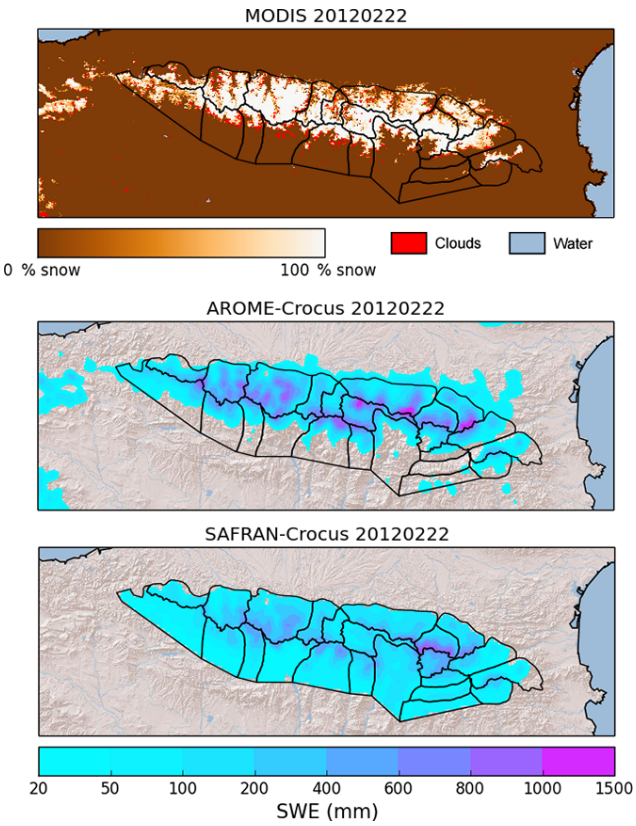

Figure 4. Top: snow cover fraction on 22 February 2012, from MO10A1 images (0.005° resolution). Bottom: SWE simulations by AROME–Crocus and SAFRAN–Crocus, same date. SAFRAN–Crocus simulations are only defined within SAFRAN massifs.

orographic shield of the Haute Bigorre first high summits leads to fewer snowfall than upstream for the same altitude (approximately 4 times less). This windward / leeward distinction within a massif is not simulated by SAFRAN, since two points at the same altitude and within the same massif get the same amount of snowfall. The difference between both forcings is marked at Esera (Spanish massif), where the orographic shield and resulting dry weather is not represented enough by SAFRAN compared to AROME. Such differences are even more marked when filtering only cumulated snowfalls occurring by N-NW flows (not shown). AROME simulations are in good agreement with the two Spanish stations, which are located at an altitude close to the model's topography. SAFRAN snowfalls are too low at the station closest to the border but in good agreement at the second Spanish station. Observations for France are in better agreement with AROME than with SAFRAN but still higher than both simulations. This may be due to the difference of altitude with the models. This study emphasises the added value of AROME dynamics, which allow us to better take into account mesoscale orographic effects.

4.1.3 Snow cover distribution

The comparison between AROME–Crocus, SAFRAN–Crocus and MODIS snow cover distribution is extended to two entire winters: 2011–2012 (characterized by an average deficit of snow) and 2012–2013 (extremely high amount of snow). Table 3 summarises two metrics (ASSD and Jaccard index) that evaluate the match of simulated and observed snow covers in different domains. AROME–Crocus scores are better than SAFRAN–Crocus scores for the whole Pyrenees (higher Jaccard index and lower ASSD for both seasons). This is also true for the Spanish, central and eastern domains, whereas scores are equivalent for France. SAFRAN–Crocus performs better in the western Pyrenees. The seasonal evolution of scores over this domain (not shown) indicates that both models have equivalent skills during the accumulation season, while SAFRAN–Crocus performs better during the melting season. This result is consistent with the results of Sect. 4.1.1: AROME–Crocus strongly overestimates snow quantities in the western Pyrenees, which results in a later presence of snow on the ground in the springtime.

Figure 6 shows the evolution of daily ASSD and Jaccard index for winter 2011–2012 over the whole Pyrenees (within SAFRAN massifs). Both scores attest that AROME–Crocus improves the representation of the spatial snow cover distribution compared to SAFRAN–Crocus until late March. SAFRAN–Crocus shows a slightly better agreement than AROME–Crocus after late March, i.e. at the beginning of the melting season due to the overestimation of snow quantities by AROME–Crocus. On 22 February 2012 (date studied in the previous section, Fig. 4), $J = 0.61$ and ASSD = 1.22 pixels for AROME–Crocus, while $J = 0.40$ and ASSD = 2.09 pixels for SAFRAN–Crocus, which quantifies the better agreement seen in Fig. 4.

4.2 Daily SD variations

4.2.1 Global scores

The STDE of daily ΔSD indicates the ability of the model to forecast (or analyse) the appropriate daily evolution of snow depth. This score was computed for AROME–Crocus and SAFRAN–Crocus. It is equal to 7 cm (and bias equal to 0 cm) for both models, with low spatial variation. STDE is slightly higher during the most snowy winters (8 cm in 2012–2013 and 2013–2014 against 6 cm in 2010–2011 and 2011–2012). This is the first complementary information to global scores that indicate that, despite an overall overestimation, AROME–Crocus gives similar results compared to SAFRAN–Crocus in terms of daily snow depth variations.

4.2.2 Categorical scores

A classification by category of the increase (accumulation) and decrease (ablation and settling) of SD gives a better view

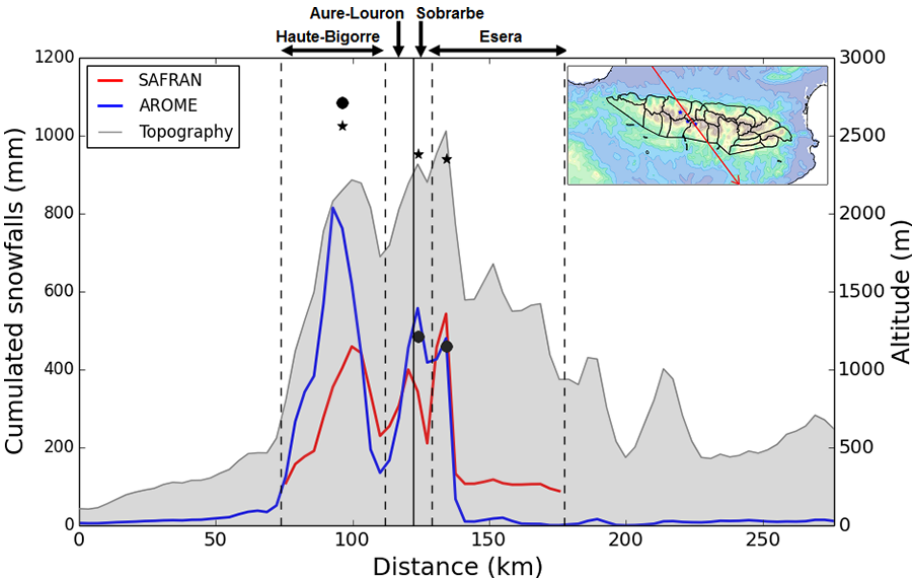

Figure 5. Cross section of cumulated snowfall from 1 October 2011 to 22 February 2012 for AROME forecasts (blue) and SAFRAN reanalysis (red), with topography plotted on the right axis in grey. Cumulated positive ΔSWE from measurements of three stations close to the transect are represented with black dots; their actual altitude is represented with black stars. The locations of the transect (red) and stations (blue stars) are given on the upper right map.

Table 3. Seasonal means of daily Jaccard index and ASSD for simulated snow cover distribution against MODIS observations in the Pyrenees for winters 2011–2012 and 2012–2013. The best scores are given in bold.

Year	Domain	*N*	Jaccard index		ASSD (pix.)	
			AROME	SAFRAN	AROME	SAFRAN
2011–2012	All	57	**0.47**	0.40	**1.34**	1.64
	France	57	0.51	**0.55**	0.91	**0.76**
	Spain	56	**0.42**	0.28	**1.27**	1.88
	West	56	0.45	**0.48**	1.34	**1.04**
	Centre	57	**0.51**	0.39	**1.08**	1.64
	East	56	**0.42**	0.31	**1.27**	1.98
2012–2013	All	39	**0.40**	0.36	**1.73**	2.00
	France	39	**0.44**	**0.44**	**1.52**	1.61
	Spain	35	**0.39**	0.32	**1.52**	2.05
	West	37	0.43	**0.45**	1.36	**1.12**
	Centre	38	**0.43**	0.37	**1.31**	1.66
	East	26	**0.42**	0.32	**1.37**	1.75

on the behaviour of the models. The categorical frequency distribution of ΔSD is plotted in Fig. 7, according to eight accumulation categories, two decrease categories and one "no variation" category [−0.2 cm, 0.2 cm[. Small daily accumulations (between 0.2 and 10 cm per day) are overrepresented by both models, while the occurrence of medium and high daily accumulations (more than 10 cm per day) is underestimated by both models. However, the frequency of medium and high accumulation events predicted by AROME–Crocus is systematically closer to the observations than SAFRAN–Crocus. There is also a clear discrepancy between both models and observations for the strong decrease category, largely underestimated by both AROME–Crocus and SAFRAN–Crocus.

In terms of quantities, the categorical sums of ΔSD (not shown) indicate that SAFRAN–Crocus strongly underestimates the high accumulation quantities. AROME–Crocus is closer to observations for these categories (particularly for the [10 cm, 20 cm[category, the main contributor to the snow accumulation). It is counterbalanced by an overestimation of small accumulation quantities, since an underestimated strong accumulation event is counted in the smaller accumulation category. The sum of all accumulation categories shows an overall underestimation of snow accumula-

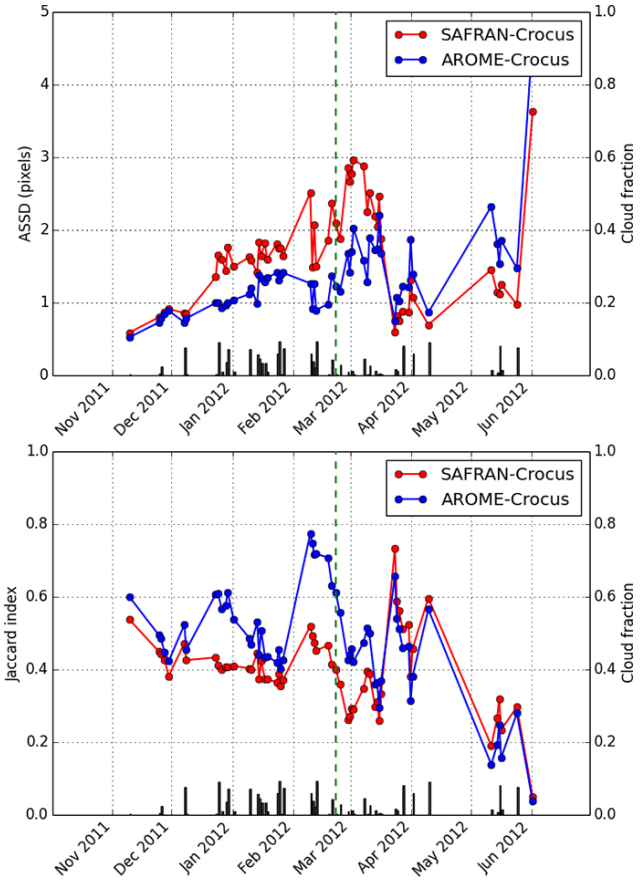

Figure 6. Daily ASSD (top) and Jaccard index (bottom), within all massifs, of AROME–Crocus vs. MODIS (blue) and SAFRAN–Crocus vs. MODIS (red), 2011–2012. Smaller ASSD and higher J mean better match with MODIS. The green line indicates 22 February 2012. The cloud fraction is represented by the black bars.

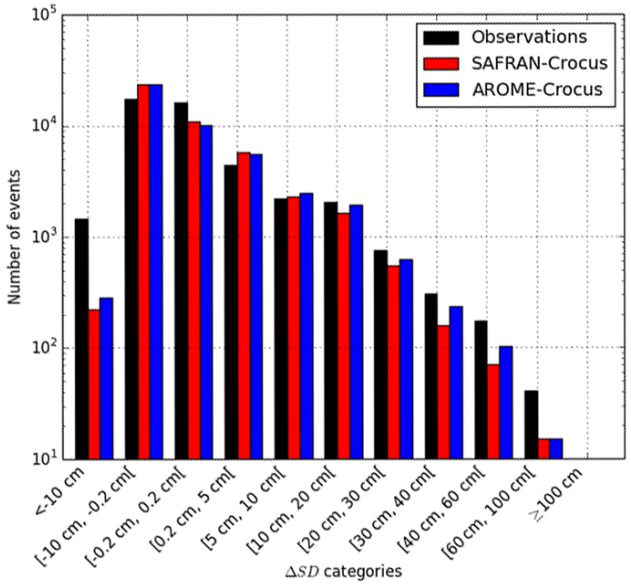

Figure 7. Categorical frequency distribution of ΔSD for observations (black), AROME–Crocus (blue) and SAFRAN–Crocus (red), at all stations, 2010–2014.

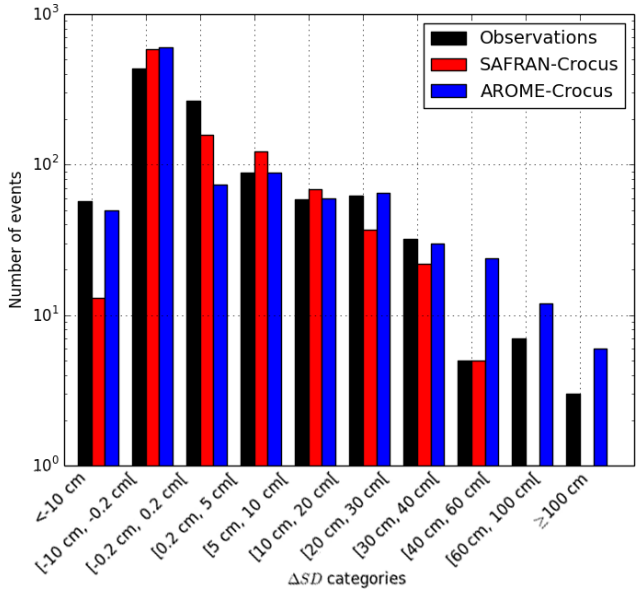

Figure 8. Categorical frequency distribution of ΔSD for observations (black), AROME–Crocus (blue) and SAFRAN–Crocus (red), at three stations near Pic d'Anie, 2010–2014.

tion by both models: the total sum of observed accumulations is 904 m, against 857 m for AROME–Crocus (−5 %), and 753 m for SAFRAN–Crocus (−17 %). The largest difference concerns the category of strong decrease, globally missed by both models. Since AROME–Crocus and SAFRAN–Crocus underestimate accumulations, the strong decrease category becomes the main contributor to the overall overestimation of snow depth: the positive bias shown in Sect. 4.1.1 is not due to an excess of snowfall but to an insufficient snow depth decrease. Total decrease quantities are more pronounced for AROME–Crocus than SAFRAN–Crocus as a logical consequence of more marked accumulations. Plotting the cumulated ΔSD by altitudinal range (under 1800 m, between 1800 and 2200, and above 2200 m) highlights a similar behaviour of both models, except for a stronger underestimation of high accumulations by SAFRAN–Crocus at the lowest altitudes (not shown).

In order to isolate the specific behaviour of AROME–Crocus in the Atlantic foothills, ΔSD categorical distribution is plotted in Fig. 8 for the three stations near Pic d'Anie, where the positive bias was found to be the highest in Sect. 4.1.1. In contrast to its general behaviour, AROME–Crocus strongly overestimates accumulations, particularly strong accumulations. At the same time, strong decreases are also underestimated, which results in a rather high positive bias.

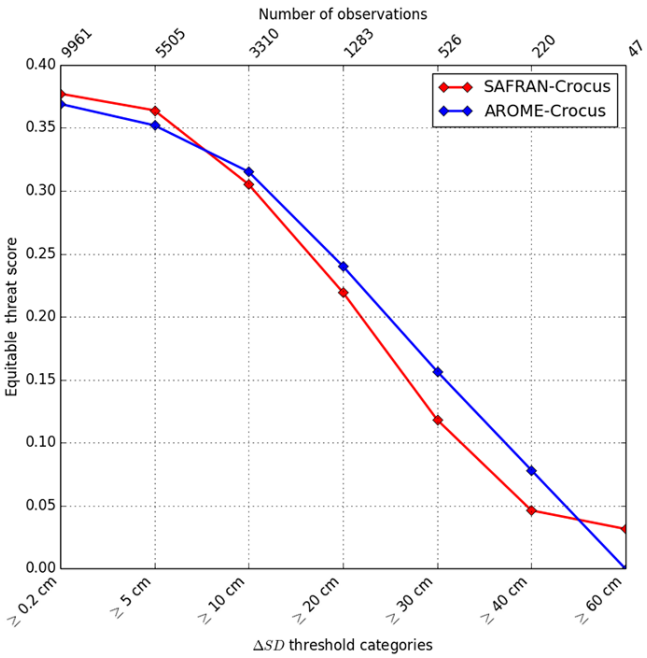

Figure 9. ETS of ΔSD threshold categories for AROME–Crocus (blue) and SAFRAN–Crocus (red), 2010–2014.

4.2.3 Study of accumulation processes and comparison to precipitations

The performance of models for daily snow accumulations is further studied thanks to the ETS, computed for threshold categories (Fig. 9). Scores are similar for AROME–Crocus and SAFRAN–Crocus. The ETS is almost 0.40 for the "all accumulations" category (more than 0.2 cm) and is under 0.10 for high accumulations (more than 40 cm). SAFRAN–Crocus has a better ETS for small accumulations, but the ETS of AROME–Crocus is better for all accumulations over 10 cm, except for extreme accumulations (more than 60 cm). However, the very small sample size for this category (47 observed events) makes impossible any reliable interpretation. A distinction by altitudinal range shows equivalent ETS for AROME–Crocus and SAFRAN–Crocus above 1800 m, and higher ETS for AROME–Crocus for medium and strong accumulations under 1800 m (not shown).

A complementary information on winter precipitation comes from the network of gauges in the French Pyrenees (red dots in Fig. 1). Daily accumulations of precipitation (rainfall plus snowfall, cumulated from 06:00 to 06:00 UTC) from the forcing models are then directly compared to precipitation gauges measurements for days with a maximum temperature of 2 °C in order to reduce the proportion of rainfall amongst precipitation. Most of these observations are assimilated in SAFRAN reanalyses, while they are not taken into account in AROME forecasts. Figure 10 shows cumulated precipitation by category for both models and observations (right) compared to cumulated ΔSD at the same stations (left). Contrary to ΔSD, AROME overestimates precipitation measured by gauges (+73 %). The optimal interpolation basis of the SAFRAN analysis system should mathematically not be biased on the assimilated observations over a long period. The slightly positive bias obtained in this study (+17 %) may be linked to the fact that some assimilated observations are not included in our evaluation dataset and/or to differences between the climatological guess and the mean precipitation amount of the 4 years under study. The strong overestimation of AROME is particularly notable for the largest amounts. The different distribution of precipitation and ΔSD for AROME, with a higher proportion of strong precipitation than of strong snow accumulations, may be due to settling effects: the stronger the snowfall, the stronger the snowpack settles under its own mass, which shifts the distribution to the left.

The overestimation of precipitation by AROME compared to precipitation gauges seems to be an apparent paradox, as we highlighted an opposite behaviour in terms of snow accumulation. This theoretical discrepancy can be explained by the quality of precipitation gauge measurements. The undercatch of solid precipitations by gauges, mainly due to wind effects on falling snowflakes trajectories, is well known and very variable. This issue is investigated by the WMO Solid Precipitation InterComparison Experiment (e.g. Wolff et al., 2015). There is no undercatch correction applied to these manual measurements, which implies that real precipitation amounts can be underestimated in the observations under windy conditions. The difference between accumulation and precipitation errors also involves modelled snow density; this issue is discussed in Sect. 4.3.

4.2.4 Study of ablation processes

A major part of model positive bias in SD is due to the underprediction of strong SD decreases. Consequently, the understanding of models biases implies a more developed study of ablation processes. Strong decreases, more than 10 cm day^{-1}, can be related to ablation processes such as melting or wind-induced erosion, which need to be studied separately. To this end, two diagnostics have been applied to identify such processes. Melting snow days (MSD) correspond to days when the snow upper layer temperature is equal to melting point at 12:00 UTC, in SAFRAN–Crocus outputs (there are no snow surface temperature measurements available). Wind-blown snow days (BSD) are identified at automatic weather stations only, where 10 m wind measurements are available. BSD correspond to days when 10 m wind speed exceeds 8 m s^{-1} during more than 10 min but no melting is diagnosed (only dry snow can be drifted). This value is based on the estimate of wind threshold for dry snow transport by Li and Pomeroy (1997). These criteria are obviously quite rough, but a comparison with snow depth plots is quite satisfactory. As an illustration, the diagnosed days are reported in Fig. 11 together with the snow depth

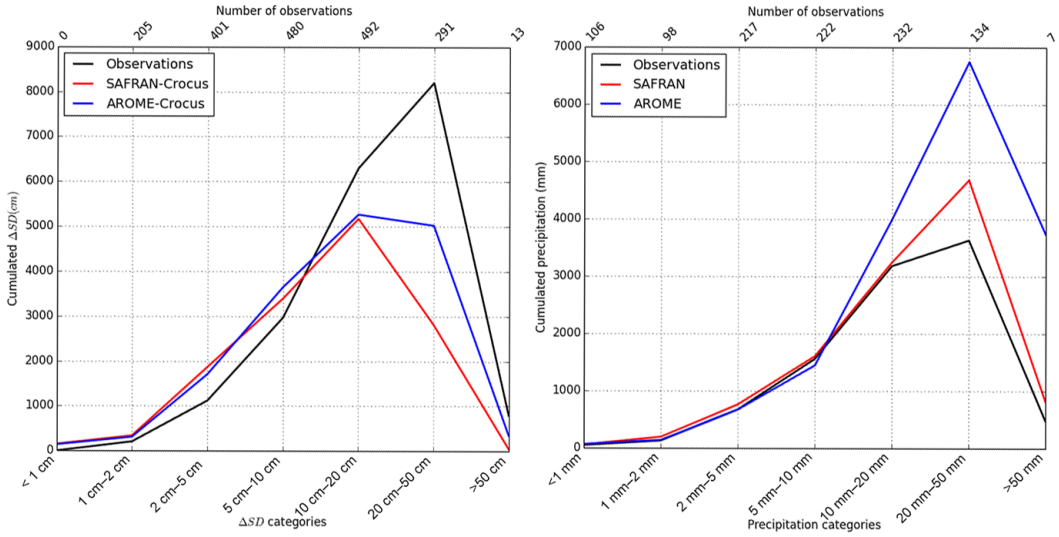

Figure 10. Cumulated ΔSD (left) and precipitation (right) for observations (black), AROME–Crocus (blue) and SAFRAN–Crocus (red), by categories, at the 28 same stations with SD and precipitation measurements during period DJFM, 2010–2014.

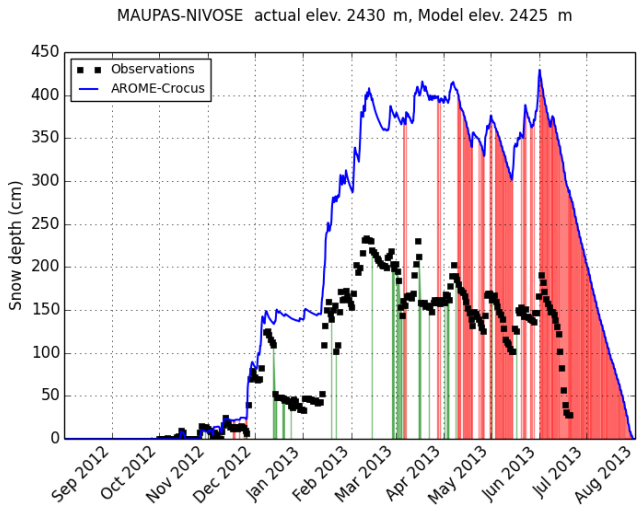

Figure 11. Snow depth simulated by AROME–Crocus (blue line) and observed (black squares) at Maupas station, 2012–2013. Wind-blown snow days are identified in green and melting snow days in red.

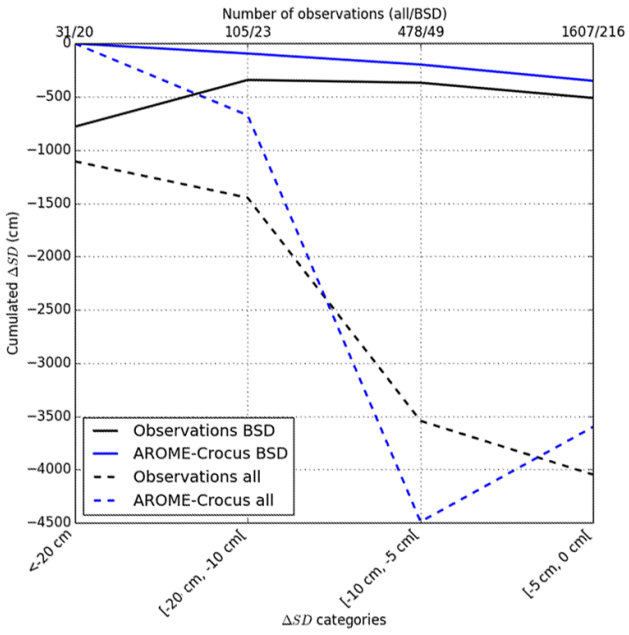

Figure 12. Cumulated ΔSD for AROME–Crocus (blue) and observations (black) by categories at seven high-altitude stations, for BSD (solid lines) and all days (dashed lines), 2010–2014.

evolution measured and simulated by AROME–Crocus, at the Maupas automatic station (massif of Luchonnais, central Pyrenees, France), where blowing snow events are known to be frequent. For instance, a good example of BSD occurred on 14 December 2012 with a 60 cm snow depth drop. MSD happen generally after April 2013 and are associated with decreasing snow depth.

To quantify the impact of wind-blown snow events on the performance of models, the cumulated ΔSD for AROME–Crocus and observations are plotted in Fig. 12, for BSD and all days, with a finer categorisation of SD decreases. This study is restricted to seven automatic stations measuring wind speed and SD (mean altitude: 2203 m.a.s.l). For observations, BSD contribute to all decreasing rates, in the strongest proportion for high decreasing rates (more than 20 cm day^{-1}). For AROME–Crocus, BSD do not contribute to the strong ablation categories but to small ablation and accumulation categories in the same proportions. Cumulated ΔSD for high decreasing rates is equal to −1106 cm for all

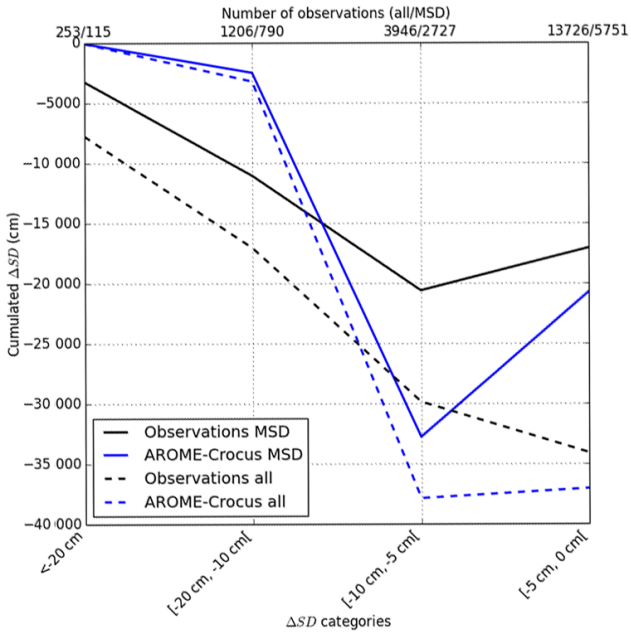

Figure 13. Cumulated ΔSD for AROME–Crocus (blue) and observations (black) by categories at all stations for MSD (solid lines) and all days (dashed lines), 2010–2014.

observations and equal to −781 cm for BSD only (excluding MSD), while it is equal to 0 cm for AROME–Crocus in both cases. It means that wind-blown snow is the main contributor (71 %) to this category, the remaining contribution coming from MSD or other processes.

Similarly, the cumulated ΔSD is plotted in Fig. 13 for MSD and all days, at all SD stations. Very strong melting (more than 20 cm day^{-1}) is sometimes observed but never predicted. Strong melting (between 10 and 20 cm day^{-1}) is much underrepresented by models, while melting of less than 10 cm day^{-1} is overrepresented. Cumulated ΔSD for high decreasing rates (more than 20 cm day^{-1}) is equal to −7741 cm for all observations and equal to −3215 cm for MSD only, while it is equal to −41 cm for AROME–Crocus in both cases. Melting snow represents 42 % of this category, the remaining contribution coming from BSD or other processes. The behaviour of SAFRAN–Crocus is similar to AROME–Crocus for BSD and MSD (not shown). The simple diagnostics for BSD and MSD may miss some wind-blown snow or melting events.

Consequently, the underestimation of strong decreasing rates comes mainly from ablation processes: on the one hand, from wind-blown snow events which are not represented by models, as they are small-scale processes; on the other hand, from an underestimation of strong snowpack melting (more than 10 cm day^{-1}). Other reasons for very high decreasing rates can be the strong settling after an intense snowfall or a rain-on-snow event, but it probably constitutes a limited part of this category.

4.3 Snow water equivalent and bulk snowpack density

20 Pyrenean stations also recorded SWE measurements from 2010–2011 to 2012–2013. Table 4 summarises the scores (bias and STDE) for SWE (upper part of the table). These stations are mainly above 2000 m a.s.l (Fig. 2) and, thus, are not representative of all SD stations of the Pyrenees. Consequently, SD scores from these stations are added at the bottom of Table 4 for an adequate comparison. While SD scores follow the tendency indicated previously (strong overestimation for AROME–Crocus, slighter overestimation for SAFRAN–Crocus), SWE scores show a lower overestimation by AROME–Crocus in relative values (+33 % for SWE, +54 % for SD, period 2010–2013) and a slight underestimation by SAFRAN–Crocus (−9 % for SWE, against +10 % for SD). The STDE is equivalent between both simulations, even slightly lower for AROME–Crocus.

It is deemed necessary to investigate further the bulk snowpack density in simulations, in order to explain the discrepancy between SWE scores and SD scores. SWE and SD measurements at the 20 automatic stations are made at the same point, which enables us to compute a bulk snowpack density: $\rho = \mathrm{SWE/SD}$ with ρ in kg m^{-3}, SWE in kg m^{-2} and SD in m. As SWE and SD measurement areas do not exactly overlap, we only consider snowpacks deeper than 20 cm to avoid problems of local heterogeneity, e.g. due to patchy snow cover during the melting season. AROME–Crocus and SAFRAN–Crocus both have a negative bias of −50 kg m^{-3} for a mean observation of 382 kg m^{-3}. The bulk snowpack density is mainly driven by the snowpack model, even when meteorological conditions are also involved. Consequently, the bias in terms of SD is necessarily higher than the bias in terms of SWE. A good simulation of SWE will lead to an overestimation of SD because of a too-low bulk snowpack density. Figure 14 shows the mean and standard deviation of simulated and observed ρ at the 20 stations, for periods of 10 days, during the 2011–2012 winter (left) and the 2012–2013 winter (right). Both winters have very different snow cover evolutions. As mentioned previously, winter 2011–2012 is characterized by a rather thin snowpack, which implies a strong variability of ρ and high bulk density during all winter. For instance, 50 cm of snow fell on bare ground at the beginning of November 2011 with no other significant occurrence during that mild month. This led to a quick settling, often associated with melting, and hence a strong densification of the thin snowpack until the beginning of December (mean observed ρ of 450 kg m^{-3}). Winter 2012–2013 was very cold and wet (Vada et al., 2013), with a very deep snowpack. A rather continuous densification of the snowpack occurred during the whole season. The negative bias of AROME–Crocus is stronger for winter 2011–2012 (−88 kg m^{-3} for a mean observation of 403 kg m^{-3}, thin and dense snowpack) than for winter 2012–2013 (−37kg m^{-3} for a mean observation of 385 kg m^{-3}, deep and less dense snowpack). Both snowpacks reached 550 to 600 kg m^{-3} (firn den-

Table 4. Scores for simulated SWE and SD against observations in 20 high-altitude automatic stations in the Pyrenees for winters 2010–2011 to 2012–2013.

SWE	Stations	*N*	Mean obs. (mm)	Bias (mm) AROME	Bias (mm) SAFRAN	STDE (mm) AROME	STDE (mm) SAFRAN
2010–2013	20	14 575	378	124	−35	272	277
2010–2011	20	4979	282	139	−5	208	179
2011–2012	20	4877	248	134	26	212	219
2012–2013	19	4719	614	96	−130	367	375
SD	**Stations**	***N***	**Mean obs. (cm)**	**Bias (cm) AROME**	**Bias (cm) SAFRAN**	**STDE (cm) AROME**	**STDE (cm) SAFRAN**
2010–2013	19	13 111	92	50	10	61	57
2010–2011	19	4405	74	53	12	50	41
2011–2012	17	4222	57	55	20	57	54
2012–2013	19	4484	142	41	−3	73	69

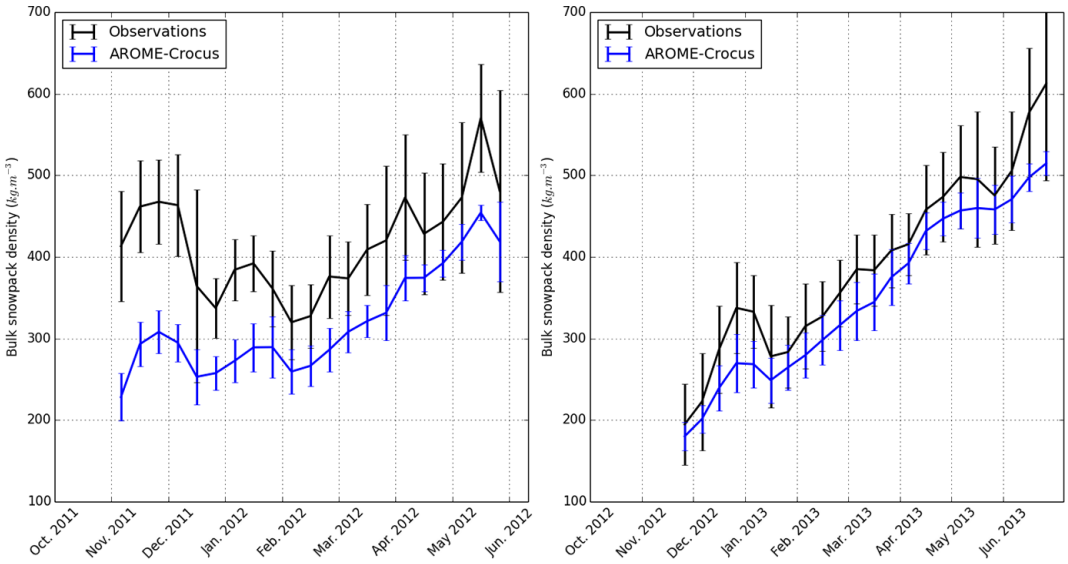

Figure 14. Bulk snowpack density during winters 2011–2012 (left) and 2012–2013 (right): mean of AROME–Crocus simulation: (blue) and observations (black) at 20 stations, for periods of 10 days. Error bars represent standard deviation.

sity) at the very end of the spring (end of May in 2012 and end of June in 2013).

A typical example of the seasonal evolution of the bulk snow density is represented in Fig. 15, at the station Les Songes (massif of Orlu, eastern Pyrenees, France), during winter 2012–2013. ρ is underestimated by AROME–Crocus during the whole season, particularly after long settling periods. Indeed, the densification slope is too low during the settling following a snowfall (increasing ρ, red arrows in Fig. 15). This is observable after every snowfall (decreasing ρ, green arrows in Fig. 15). For instance, fresh snow falls at the beginning of December 2012, with an adequate simulation of ρ until then; the process of settling and densification of the snowpack occurs during the whole month of December reaching 350 kg m^{-3} in observations, while the densification slope is much lower in simulations, reaching less than 300 kg m^{-3}.

5 Discussion and conclusion

A more accurate description of the snow cover variability in mountainous terrain is necessary for many applications including mountain hydrology or avalanche hazard forecasting. In this paper, we have addressed the potential of the kilometre-scale NWP model AROME used as atmospheric forcing for distributed snowpack simulations in the Pyrenees. The simulations were carried out with the snowpack model Crocus at a 2.5 km grid spacing, during four contrasted winters, from August 2010 to August 2014. They were evalu-

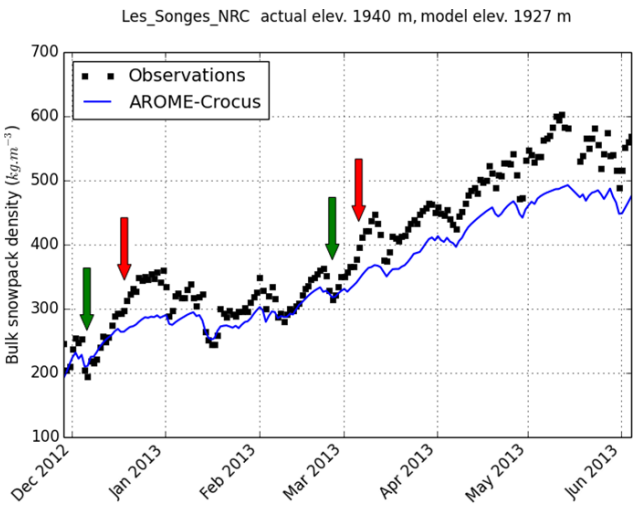

Figure 15. Bulk snowpack density observed (black) and simulated by AROME–Crocus (blue) at station Les Songes, winter 2012–2013. Green arrows indicate two examples of snowfalls and red arrows indicate two examples of settling period.

ated through a comparison to simulations driven by the analysis system SAFRAN and to ground-based measurements of snow depth, snow water equivalent and precipitation across the whole mountainous chain, as well as MODIS images of snow cover fraction. A global verification of snow depth simulation with 83 stations exhibited an overestimation in both simulations, with a higher positive bias for AROME–Crocus than SAFRAN–Crocus. In terms of SWE (20 stations), the overestimation was less marked for AROME–Crocus and turned out to be an underestimation for SAFRAN–Crocus. Compared to the evaluation performed by Vionnet et al. (2016) in the French Alps, the overestimation by AROME–Crocus is stronger in the Pyrenees (+55 cm against +40 cm in the Alps) and, to a lesser extent, by SAFRAN–Crocus too (+22 cm against +17 cm in the Alps). This overestimation may originate from the immediate vicinity and influence of the Atlantic Ocean and the Mediterranean Sea. However, for a longer time period, SAFRAN–Crocus does not exhibit such a bias over the French Pyrenees (Lafaysse et al., 2013), and the results may be specific to the studied seasons. The lowest biases were found in the eastern part of the Pyrenees, which is also the driest, a result similar to that of Vionnet et al. (2016), who highlighted a lower overestimation in the southern Alps. The highest biases were found in the western Pyrenees, where precipitations from the Atlantic Ocean come first and in the greatest quantity.

AROME–Crocus exhibits a better snow spatial distribution than SAFRAN–Crocus with respect to MODIS images of snow cover fraction. Similarity scores highlighted a better agreement of snow-covered areas for AROME–Crocus, for two winters in most domains, except in the western Pyrenees where AROME snowfalls are too large. The added value of AROME–Crocus to represent the spatial variability of the snowpack within each massif was particularly emphasized on winter 2011–2012. AROME captures mesoscale orographic effects (enhanced precipitation on the upwind side of mountains, as shown in Fig. 5), thus enabling a more adequate distribution of the snow cover compared to SAFRAN–Crocus. Vionnet et al. (2016) showed this high variability within Alpine massifs in terms of seasonal snowfall. The dynamical behaviour of AROME, compared to SAFRAN, is of particular interest in a relatively narrow chain such as the Pyrenees, where orographic blocking and foehn effects are very frequent, creating strong climatic and snowpack heterogeneities. Nevertheless, the orographic blocking was shown to be excessive for mountains closest to the Atlantic Ocean, which is probably due either to an excessive vertical updraft of the disturbed oceanic flows on the first steep slopes or to an excessive model reactivity to these updrafts.

The study of daily SD and SWE variations enables a more detailed understanding of the scores of models. We indeed show that the global overestimation of SD and SWE is not the consequence of overestimated snowfall (except in the Atlantic foothills). Snow accumulation, and especially strong accumulation, are underestimated by both AROME–Crocus and SAFRAN–Crocus, with AROME–Crocus performing best. These results are in total agreement with the study of Schirmer and Jamieson (2015), using GEM-LAM (2.5 km resolution NWP model, equivalent to AROME, Erfani et al., 2005) and GEM15 (15 km resolution NWP model, equivalent to ARPEGE, Mailhot et al., 2006) as atmospheric forcing to SNOWPACK (detailed snowpack model, equivalent to Crocus; Bartelt and Lehning, 2002). They showed the same underestimation of strong accumulations, less marked for the high-resolution forcing. The ETS of GEM-LAM/SNOWPACK for ΔSD accumulation threshold categories is very close to the ETS shown here for AROME–Crocus.

The comparison with precipitation gauges did not confirm the underestimation of snow accumulations since precipitation seemed to be overestimated by AROME, but this paradox can be explained by the uncorrected undercatch of winter precipitation. The assimilation of these data in SAFRAN precipitation analysis tends to reduce them excessively and subsequently greatly reduce snow accumulations in SAFRAN–Crocus. The problematic assimilation of precipitation gauge measurements in mountainous terrain is also underlined by Schirmer and Jamieson (2015) for the Canadian Precipitation Analysis system (CaPA) (Mahfouf et al., 2007). This study thus tends to substantiate the idea that variations of SD and SWE measured on the ground could replace precipitation gauges in precipitation analyses in mountainous terrain, as evoked by Schirmer and Jamieson (2015). Magnusson et al. (2014) also showed that point SWE data assimilation could improve distributed snow cover model simulations.

The underestimation of snow accumulation is counterbalanced by an underestimation of the intensity of ablation pro-

cesses. We first showed that wind-induced erosion of the snowpack constituted the major cause of the underestimation of strong ablations at seven high-altitude stations. This small-scale process cannot be captured by a kilometric simulation of the snowpack, since snow redistribution by wind occurs very likely within each grid cell. However, the computation of SD and SWE scores is affected by the occurrence of wind-induced snow transport at stations. The impact of blowing snow could not be estimated at all stations. It is probably less significant at lower altitudes. Secondly, we showed that the intensity of strong melting is underestimated. This process has several sources which need to be further explored. Candidates for possible sources are the physical description of melting within the snowpack model, the incoming shortwave and long-wave radiations in the atmospheric forcing affecting the snowpack surface energy balance, the formulation of turbulent fluxes. Furthermore, this result is in contradiction with the evaluation of the Crocus model forced by in situ meteorological measurements (Brun et al., 1992; Vionnet et al., 2012), where such a bias has never been noticed. It will be essential to refine the evaluation of the snowpack model in such conditions using the modus operandi described in this paper. Finally, a simultaneous study of the evolution of SWE and SD gave the opportunity to evaluate the simulated bulk snowpack density. A global underestimation was shown for AROME–Crocus, supporting the hypothesis of an insufficient settling of the snowpack after a snowfall in Crocus. This hypothesis is consistent with previous simulations at the Col de Porte station in the Alps (not shown). Consequently, all processes contributing to the decrease of the snow depth are underestimated, in a stronger proportion than for accumulations, which leads to a global overestimation of snow depths through a smoothing of extreme variations. These opposite biases artificially imply a smaller bias for SAFRAN–Crocus than for AROME–Crocus. The underestimation of the intensity of daily variations also implies daily variations of the bias, hence a high dispersion around the mean bias, which partly explains a high STDE. This daily-scale study thus highlights the limitations of global scores (bias, root mean square error, STDE) for a physical quantity like snow depth, which depends on several physical processes. Another limitation is the cumulative error during the winter season. The representativeness of stations, which are influenced by local phenomena, may also be questioned (Grünewald and Lehning, 2015), although the large sample of stations, with a large spatial and altitudinal distribution, may reduce the impact of such issues in the present study.

Several limitations also have to be tackled concerning the daily variations of SD and SWE. Data series need to be processed very carefully, since one odd value in the observations would have a double impact in terms of daily variations. Moreover, the daily increase of the snow depth includes not only fresh snowfall but also its own settling and the settling of the underlying layers during 1 day. This phenomenon tends to reduce the estimated snow accumulation. Following Fischer (2011), a time interval of 6 hours would be more appropriate, but the availability of measurements only made it possible for the automatic stations. ΔSWE measurements enable to put the issue of snow settling aside, since it does not affect the snowpack mass. However, SWE measurements by cosmic ray snow gauges are associated with noise due to atmospheric conditions (Gottardi et al., 2013) and thus requires a 24 h median smoothing, which subsequently limits the accuracy of ΔSWE values to $\pm 10\,\%$. Finally, daily variations of snowpack depth or mass are strongly impacted by wind-blown snow events, as shown in Fig. 12: beyond the inherent information about such events, using measurements of snow on the ground to derive snowfall quantities would require a correction by additional information from snowdrift measurements, as suggested by Fischer (2011).

These results underline the relevance of AROME–Crocus forecasts to provide high-resolution spatial patterns of the snowpack in the Pyrenees, while Vionnet et al. (2016) got similar results in the French Alps. What remains is to use this potential in the assimilation of observations in mountainous terrain so as to implement a spatially distributed meteorological analysis system, which would substantially improve the atmospheric forcing as was the case at massif scale with SAFRAN (Durand et al., 1993). Indeed, most of the uncertainties of a snowpack simulation come from the atmospheric forcing (Raleigh et al., 2015). To deal with that, the use of complementary observations in complex terrain is necessary, with a particular emphasis on precipitation. For instance, Birman et al. (2016) recently developed a new precipitation analysis system, combining a priori information from AROME with ground-based and radar observations. Satellite cloud masks could also be used to improve incoming radiations (e.g. Hinkelman et al., 2015), and new polarimetric radar products could help to determine the snow / rain limit (e.g. Augros et al., 2015). The development of higher-resolution versions of AROME or the use of downscaling methods on the meteorological forcing (Vionnet et al., 2015) would enable sub-kilometric snowpack simulations to take into account effects of slope and aspect on incoming radiations. Additionally, observations can also be assimilated directly within the snowpack model, e.g. as done by Charrois et al. (2016) for optical reflectances in the Crocus model. Finally, as all errors cannot be eliminated, the potential of using ensemble high-resolution forecasts should also be explored. The benefit in forecasting extreme hydrological events has been demonstrated (Vié et al., 2011), and Vernay et al. (2015) illustrated the advantage of using ensemble forecasting for avalanche hazard assessment.

Significant benefits can also be derived from AROME short-range forecasts: further studies at shorter timescales would shed light on AROME potential for snowpack evolution forecast for high impact events, like intense snowfall triggering off avalanches, rain-on-snow events or ice layer formation.

6 Data availability

SD and meteorological variables measurements from Météo France stations and AROME forecasts in real time are publicly available at https://donneespubliques.meteofrance.fr. All AROME forecast archives are available on request from the same website for a data provision fee, but the variables, domain and period used in this study are available for research purposes on request from the authors. SD and SWE measurements were also provided by Electricité De France, Confederación Hidrográfica del Ebro, Servei Meteorològic de Catalunya, Centre d'Etudes Spatiales de la Biosphère and Instituto Pirenaico de Ecología: they should be contacted directly for data access. MODIS fractional snow cover images are available at https://nsidc.org/data/mod10a1. SAFRAN analyses are available for research purposes on request from the authors.

Acknowledgements. The authors acknowledge Electricité De France and Confederación Hidrográfica del Ebro for providing snow water equivalent and snow depth measurements from their Pyrenean automatic stations network, Servei Meteorològic de Catalunya for providing snow depth measurements from the automatic weather stations network of Catalunya, S. Gascoin (CESBIO) for providing snow depth measurements from automatic station Bassies, and J. Revuelto (Instituto Pirenaico de Ecología) for providing snow depth measurements from automatic station Izas. We particularly thank E. Bazile (CNRM), F. Gottardi (EDF-DTG), S. Morin (CNRM/CEN), R. Mott (WSL-SLF) and B. Vincendon (CNRM) for help and discussions and Jean-Antoine Maziejewski (Météo France) for English editing. The authors are also grateful to R. Essery and the anonymous reviewer for their detailed comments on the manuscript.

Edited by: R. Brown
Reviewed by: R. L. H. Essery and one anonymous referee

References

Anderton, S. P., White, S. M., and Alvera, B.: Micro-scale spatial variability and the timing of snow melt runoff in a high mountain catchment, J. Hydrol., 268, 158–176, doi:10.1016/S0022-1694(02)00179-8, 2002.

Augros, C., Caumont, O., Ducrocq, V., Gaussiat, N., and Tabary, P.: Comparisons between S-, C- and X-band polarimetric radar observations and convective-scale simulations of the HyMeX first special observing period, Q. J. R. Meteorol. Soc., doi:10.1002/qj.2572, 2015.

Bartelt, P. and Lehning, M.: A physical SNOWPACK model for the Swiss avalanche warning: Part I: numerical model, Cold Reg. Sci. Technol., 35, 123–145, doi:10.1016/S0165-232X(02)00074-5, 2002.

Bélair, S., Roch, M., Leduc, A.-M., Vaillancourt, P. A., Laroche, S., and Mailhot, J.: Medium-Range Quantitative Precipitation Forecasts from Canada's New 33-km Deterministic Global Operational System, Weather Forecast., 24, 690–708, doi:10.1175/2008WAF2222175.1, 2009.

Bellaire, S., Jamieson, J. B., and Fierz, C.: Forcing the snow-cover model SNOWPACK with forecasted weather data, The Cryosphere, 5, 1115–1125, doi:10.5194/tc-5-1115-2011, 2011.

Bellaire, S., Jamieson, J. B., and Fierz, C.: Corrigendum to "Forcing the snow-cover model SNOWPACK with forecasted weather data" published in The Cryosphere, 5, 1115–1125, 2011, The Cryosphere, 7, 511–513, doi:10.5194/tc-7-511-2013, 2013.

Bellaire, S., Katurji, M., Schulmann, T., and Hobman, A.: Towards a High-Resolution Operational Forecasting Tool for the Southern Alps - New Zealand, in: Proceedings of the International Snow Science Workshop, Banff, Canada, 388–393, doi:10.13140/2.1.3376.8640, 2014.

Birman, C., Karbou, F., Mahfouf, J., Lafaysse, M., Durand, Y., Giraud, G., Mérindol, L., and Hermozo, L.: Precipitation analysis over the French Alps using a variational approach and study of potential added value of ground based Radar observations, J. Hydrometeor., 2016.

Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, Q. J. R. Meteorol. Soc., doi:10.1002/qj.2822, 2016.

Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13 – 22, 1992.

Bubnová, R., Hello, G., Bénard, P., and Geleyn, J.-F.: Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system, Mon. Weather Rev., 123, 515–535, doi:10.1175/1520-0493(1995)123<0515:IOTFEE>2.0.CO;2, 1995.

Buisan, S. T., Saz, M. A., and López-Moreno, J. I.: Spatial and temporal variability of winter snow and precipitation days in the western and central Spanish Pyrenees, Int. J. Climatol., 35, 259–274, doi:10.1002/joc.3978, 2015.

Charrois, L., Cosme, E., Dumont, M., Lafaysse, M., Morin, S., Libois, Q., and Picard, G.: On the assimilation of optical reflectances and snow depth observations into a detailed snowpack model, The Cryosphere, 10, 1021–1038, doi:10.5194/tc-10-1021-2016, 2016.

Dombrowski-Etchevers, I., Quéno, L., Karbou, F., Ribaud, J.-F., and Durand, Y.: Test and Potentialities of a New Numerical Weather Forecasting Non-Hydrostatic Model for Hydrology and Snow-cover Simulations, in: Proceedings of the International Snow Science Workshop, Grenoble – Chamonix Mont-Blanc, France, 1309–1314, 2013.

Dubuisson, M.-P. and Jain, A. K.: A Modified Hausdorff Distance for Object Matching, in: Proceedings of the International Conference on Pattern Recognition, Jerusalem, Israel, 566–568, 1994.

Durand, Y., Brun, E., Mérindol, L., Guyomarc'h, G., Lesaffre, B., and Martin, E.: A meteorological estimation of relevant parameters for snow models, Ann. Glaciol., 18, 65–71, 1993.

Durand, Y., Giraud, G., Brun, E., Mérindol, L., and Martin, E.: A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting, J. Glaciol., 45, 469–484, 1999.

Durand, Y., Giraud, G., Laternser, M., Etchevers, P., Mérindol, L., and Lesaffre, B.: Reanalysis of 47 Years of Climate in the French Alps (1958–2005): Climatology and Trends

for Snow Cover, J. Appl. Meteor. Climatol., 48, 2487–2512, doi:10.1175/2009JAMC1810.1, 2009a.

Durand, Y., Giraud, G., Laternser, M., Etchevers, P., Mérindol, L., and Lesaffre, B.: Reanalysis of 44 Yr of Climate in the French Alps (1958–2002): Methodology, Model Validation, Climatology, and Trends for Air Temperature and Precipitation, J. Appl. Meteor. Climatol., 48, 429–449, doi:10.1175/2008JAMC1808.1, 2009b.

Durand, Y., Giraud, G., Goetz, D., Maris, M., and Payen, V.: Modeled Snow Cover in Pyrenees Mountains and Cross-Comparisons Between Remote-Sensed and Land-Based Observation Data, in: Proceedings of the International Snow Science Workshop, Anchorage, Alaska, 998–1004, 2012.

Erfani, A., Mailhot, J., Gravel, S., Desgagné, M., King, P., Sills, D., McLennan, N., and Jacob, D.: The high resolution limited area version of the Global Environmental Multiscale model (GEM-LAM) and its potential operational applications, in: 11th Conference on Mesoscale Processes, American Meteorological Society, Albuquerque, NM, USA, 2005.

Essery, R., Morin, S., Lejeune, Y., and Menard, C. B.: A comparison of 1701 snow models using observations from an alpine site, Adv. Water Resour., 55, 131–148, doi:10.1016/j.advwatres.2012.07.013, 2013.

FAO/IIASA/ISRIC/ISS-CAS/JRC: Harmonized World Soil Database (version 1.2), Tech. rep., FAO, Rome, Italy and IIASA, Laxenburg, Austria, 2012.

Fischer, A. P.: The Measurement Factors in Estimating Snowfall Derived from Snow Cover Surfaces Using Acoustic Snow Depth Sensors, J. Appl. Meteor. Climatol., 50, 681–699, doi:10.1175/2010JAMC2408.1, 2011.

Gascoin, S., Hagolle, O., Huc, M., Jarlan, L., Dejoux, J.-F., Szczypta, C., Marti, R., and Sánchez, R.: A snow cover climatology for the Pyrenees from MODIS snow products, Hydrol. Earth Syst. Sci., 19, 2337–2351, doi:10.5194/hess-19-2337-2015, 2015.

Gottardi, F., Paquet, E., Carrier, P., Laval, M.-T., Gailhard, J., and Garcon, R.: A decade of snow water equivalent monitoring in the French Mountain ranges, in: Proceedings of the International Snow Science Workshop Grenoble, Chamonix Mont-Blanc, 2013, 7–11 October, Grenoble, France, 926–930, 2013.

Grünewald, T. and Lehning, M.: Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements at the small catchment scale, Hydrol. Process., 29, 1717–1728, doi:10.1002/hyp.10295, 2015.

Hinkelman, L. M., Lapo, K. E., Cristea, N. C., and Lundquist, J. D.: Using CERES SYN Surface Irradiance Data as Forcing for Snowmelt Simulation in Complex Terrain, J. Hydrometeorol., 16, 2133–2152, doi:10.1175/JHM-D-14-0179.1, 2015.

Horton, S., Schirmer, M., and Jamieson, B.: Meteorological, elevation, and slope effects on surface hoar formation, The Cryosphere, 9, 1523–1533, doi:10.5194/tc-9-1523-2015, 2015.

Jonas, T., Geiger, F., and Jenny, H.: Mortality pattern of the Alpine chamois: the influence of snow-meteorological factors, Ann. Glaciol., 49, 56–62, doi:10.3189/172756408787814735, 2008a.

Jonas, T., Rixen, C., Sturm, M., and Stoeckli, V.: How alpine plant growth is linked to snow cover and climate variability, J. Geophys. Res., 113, G03013, doi:10.1029/2007JG000680, 2008b.

Klein, A. G. and Stroeve, J.: Development and validation of a snow albedo algorithm for the MODIS instrument, Ann. Glaciol., 34, 45–52, doi:10.3189/172756402781817662, 2002.

Lafaysse, M., Morin, S., Coleou, C., Vernay, M., Serca, D., Besson, F., Willemet, J.-M., Giraud, G., and Durand, Y.: Towards a new chain of models for avalanche hazard forecasting in French mountain ranges, including low altitude mountains, in: Proceedings of International Snow Science Workshop Grenoble–Chamonix Mont-Blanc, 162–166, CEN, http://arc.lib.montana.edu/snow-science/objects/ISSW13_paper_O1-02.pdf (last access: 18 July 2016), 2013.

Lafore, J., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., Fischer, C., Hereil, P., Mascart, P., Pinty, J., Redelsperger, J. L., Richard, E., and Vila-Guerau de Arellano, J.: The Meso-NH Atmospheric Simulation System. Part I: adiabatic formulation and control simulations, Ann. Geophysicae, 16, 90–109, doi:10.1007/s00585-997-0090-6, 1998.

Lehning, M., Löwe, H., Ryser, M., and Raderschall, N.: Inhomogeneous precipitation distribution and snow transport in steep terrain, Water Resour. Res., 44, W07404, doi:10.1029/2007WR006545, 2008.

Li, L. and Pomeroy, J. W.: Estimates of Threshold Wind Speeds for Snow Transport Using Meteorological Data, J. Appl. Meteor., 36, 205–213, doi:10.1175/1520-0450(1997)036<0205:EOTWSF>2.0.CO;2, 1997.

López-Moreno, J. I., Goyette, S., and Beniston, M.: Impact of climate change on snowpack in the Pyrenees: Horizontal spatial variability and vertical gradients, J. Hydrol., 374, 384–396, doi:10.1016/j.jhydrol.2009.06.049, 2009.

Magnusson, J., Gustafsson, D., Hüsler, F., and Jonas, T.: Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods, Water Resour. Res., 50, 7816–7835, doi:10.1002/2014WR015302, 2014.

Mahfouf, J.-F., Brasnett, B., and Gagnon, S.: A Canadian precipitation analysis (CaPA) project: Description and preliminary results, Atmos.-Ocean, 45, 1–17, doi:10.3137/ao.v450101, 2007.

Mailhot, J., Bélair, S., Lefaivre, L., Bilodeau, B., Desgagné, M., Girard, C., Glazer, A., Leduc, A.-M., Méthot, A., Patoine, A., Plante, A., Rahill, A., Robinson, T., Talbot, D., Tremblay, A., Vaillancourt, P., Zadra, A., and Qaddouri, A.: The 15-km version of the Canadian regional forecast system, Atmos.-Ocean, 44, 133–149, doi:10.3137/ao.440202, 2006.

Maris, M., Giraud, G., Durand, Y., Navarre, J.-P., and Mérindol, L.: Results of 50 years of climate reanalysis in the French Pyrenees (1958-2008) using the SAFRAN and CROCUS models, in: Proceedings of the International Snow Science Workshop, Davos, Switzerland, 219–223, 2009.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960, doi:10.5194/gmd-6-929-2013, 2013.

Mott, R., Schirmer, M., Bavay, M., Grünewald, T., and Lehning, M.: Understanding snow-transport processes shaping the mountain snow-cover, The Cryosphere, 4, 545–559, doi:10.5194/tc-4-545-2010, 2010.

Mott, R., Scipión, D., Schneebeli, M., Dawes, N., Berne, A., and Lehning, M.: Orographic effects on snow deposition patterns in mountainous terrain, J. Geophys. Res. Atmos., 119, 1419–1439, doi:10.1002/2013JD019880, 2014.

Nurmi, P.: Recommendations on the verification of local weather forecasts, Tech. Rep. 430, European Centre for Medium-Range Weather Forecasts, 2003.

Pomeroy, J. W. and Gray, D. M.: Snowcover accumulation, relocation and management, National Hydrology Research Institute Science Report No. 7, NHRI Environment Canada, Saskatoon, 1995.

Raleigh, M. S., Lundquist, J. D., and Clark, M. P.: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework, Hydrol. Earth Syst. Sci., 19, 3153–3179, doi:10.5194/hess-19-3153-2015, 2015.

Schirmer, M. and Jamieson, B.: Verification of analysed and forecasted winter precipitation in complex terrain, The Cryosphere, 9, 587–601, doi:10.5194/tc-9-587-2015, 2015.

Schweizer, J., Jamieson, J. B., and Schneebeli, M.: Snow avalanche formation, Rev. Geophys., 41, 1016, doi:10.1029/2002RG000123, 2003.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective scale operational model, Mon. Weather Rev., 129, 976–991, doi:10.1175/2010MWR3425.1, 2011.

Sirguey, P.: Monitoring Snow Cover and Modelling Catchment Discharge With Remote Sensing in the Upper Waitaki Basin, New Zealand, Ph.D. thesis, University of Otago, Dunedin, New Zealand, 2009.

Vada, J. A., Rodriguez-Marcos, J., Buisan, S., and Ambrosio, I. S.: Climatological comparison of 2011–2012 and 2012–2013 snow seasons in Central and Western Spanish Pyrenees and its relationship with the North Atlantic Oscillation (NAO), in: International Snow Science Workshop Grenoble, Chamonix Mont-Blanc, 2013.

Vernay, M., Lafaysse, M., Mérindol, L., Giraud, G., and Morin, S.: Ensemble forecasting of snowpack conditions and avalanche hazard, Cold Reg. Sci. Technol., 120, 251–262, doi:10.1016/j.coldregions.2015.04.010, 2015.

Vié, B., Nuissier, O., and Ducrocq, V.: Cloud-Resolving Ensemble Simulations of Mediterranean Heavy Precipitating Events: Uncertainty on Initial Conditions and Lateral Boundary Conditions, Mon. Weather Rev., 139, 403–423, doi:10.1175/2010MWR3487.1, 2011.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geosci. Model Dev., 5, 773–791, doi:10.5194/gmd-5-773-2012, 2012.

Vionnet, V., Martin, E., Masson, V., Guyomarc'h, G., Naaim-Bouvet, F., Prokop, A., Durand, Y., and Lac, C.: Simulation of wind-induced snow transport and sublimation in alpine terrain using a fully coupled snowpack/atmosphere model, The Cryosphere, 8, 395–415, doi:10.5194/tc-8-395-2014, 2014.

Vionnet, V., Bélair, S. ans Girard, C., and Plante, A.: Wintertime Subkilometer Numerical Forecasts of Near-Surface Variables in the Canadian Rocky Mountains, Mon. Weather Rev., 143, 666–686, doi:10.1175/MWR-D-14-00128.1, 2015.

Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Quéno, L., Seity, Y., and Bazile, E.: Numerical weather forecasts at kilometer scale in the French Alps: evaluation and applications for snowpack modelling, J. Hydrometeor., accepted, 2016.

Wolff, M. A., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T., and Brækkan, R.: Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study, Hydrol. Earth Syst. Sci., 19, 951–967, doi:10.5194/hess-19-951-2015, 2015.

Xue, M., Droegemeier, K. K., and Wong, V.: The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification, Meteorol. Atmos. Phys., 75, 161–193, doi:10.1007/s007030070003, 2000.