# Spatial consistency and bias in avalanche forecasts - a case study in the European Alps

**Frank Techel[1,2], Christoph Mitterer[5], Elisabetta Ceaglio[3], Cécile Coléou[4], Samuel Morin[6], Francesca Rastelli[7], and Ross S. Purves[2]**

[1]WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland
[2]Department of Geography, University of Zurich Zurich, Switzerland
[3]Fondazione Montagna sicura, Ufficio neve e valanghe, Regione Autonoma Valle d'Aosta, Italy
[4]Météo France, Direction des Opérations pour la Prévision, Cellule Montagne Nivologie, Grenoble, France
[5]Lawinenwarndienst Tirol, Abteilung Zivil- und Katastrophenschutz, Innsbruck, Austria
[6]Météo France - CNRS, CNRM UMR 3589, Centre d'Études de la Neige, Grenoble, France
[7]Meteomont Carabinieri, Bormio, Italy

*Correspondence to:* Frank Techel (techel@slf.ch)

**Abstract.**

In the European Alps, the public is provided with regional avalanche forecasts, issued by about 30 forecast centers throughout the winter, covering a spatially contiguous area. A key element in these forecasts is the communication of avalanche danger according to the five-level, ordinal European Avalanche Danger Scale (EADS). Consistency in the application of the avalanche danger levels by the individual forecast centers is essential to avoid misunderstandings or misinterpretations by users, particularly those utilizing bulletins issued by different forecast centers. As the quality of avalanche forecasts is difficult to verify, due to the categorical nature of the EADS, we investigated forecast goodness by focusing on spatial consistency and bias exploring real forecast danger levels from four winter seasons (477 forecast days). We describe the operational constraints associated with the production and communication of the avalanche bulletins, and we propose a methodology to quantitatively explore spatial consistency and bias. We note that the forecast danger level agreed significantly less often when compared across national and forecast center boundaries (about 60%), as compared to within forecast center boundaries (about 90%). Furthermore, several forecast centers showed significant systematic differences towards using more frequently lower (or higher) danger levels than their neighbors. Discrepancies seemed to be greatest when analyzing the proportion of forecasts with danger level *4-High* and *5-Very High*. The size of the warning regions, the smallest geographically clearly specified areas underlying the forecast products, differed considerably between forecast centers. Region size also had a significant impact on all summary statistics and is a key parameter influencing the issued danger level, but also limits the communication of spatial variations in the danger level. Operational constraints in the production and communication of avalanche forecasts and variation in the ways the EADS is interpreted locally may contribute to inconsistencies, and may be potential sources for misinterpretation by forecast users. All these issues highlight the need to further harmonize the forecast production process and the way avalanche hazard is communicated to increase consistency, and hence facilitate cross-border forecast interpretation by traveling users.

## 1 Introduction

In the European Alps, public forecasts of avalanche hazard are provided throughout the winter. These forecasts - also called advisories, warnings, or bulletins[1] - provide information about the current and forecast snow and avalanche conditions in a specific region. In contrast to local avalanche forecasting, e.g. for a transportation corridor or ski area, a regional forecast does not provide informa-

[1]we use these terms synonymously

tion regarding individual slopes or specific endangered objects.

One of the key consumer groups are those undertaking recreational activities, such as off-piste riding and backcountry touring in unsecured terrain. The importance of clearly communicating to this group is underlined firstly by avalanche accident statistics - with on average 100 fatalities each winter in the Alps (Techel et al., 2016), most of whom died during recreational activities. Secondly, very large numbers of individuals recreate in unsecured winter terrain, with for example Winkler et al. (2016) reporting that more than two million winter backcountry touring days were undertaken in 2013 in Switzerland alone. An additional consumer group are local, regional and national risk management authorities, who base risk reduction strategies such as avalanche control measures, road closures, evacuation procedures etc. in part on information provided in regional avalanche forecasts.

In all Alpine countries (Fig. 1), forecasts are disseminated throughout the entire winter, for individual warning regions, together forming a spatially contiguous area covering the entire Alpine region. Furthermore, in all of these countries the European Avalanche Danger Scale (EADS; EAWS, 2018), introduced in 1993 (SLF, 1993), is used in the production and communication of forecasts (EAWS, 2017c).

The EADS is an ordinal, five-level scale, focusing on avalanche hazard, with categorical descriptions for each danger level describing snowpack (in)stability, avalanche release probability, expected size and number of avalanches and the likely distribution of triggering spots (Tab. 1). The EADS describes situations with spontaneous avalanches but also conditions where an additional load - such as a person skiing a slope - can trigger an avalanche. These categorical descriptions of each danger level aim to inform users on the nature of avalanche hazard at hand. However, individual danger levels capture a wide range of differing avalanche conditions (e.g. EAWS, 2005; Lazar et al., 2016; EAWS, 2017a; Statham et al., 2018a), and therefore, in isolation, are too basic to be used as a standalone decision making tool (e.g. Météo France, 2012). Additionally, and in order to describe the avalanche hazard in more detail and to provide better advice to the end users on how to manage these hazards, the EAWS introduced a set of five typical avalanche problems (EAWS, 2017d). Nonetheless, the EADS provides a consistent way of communicating avalanche hazard. Furthermore, the EADS often forms an important input into basic avalanche education on planning, or decision making heuristics as practiced by many recreationists (e.g. Munter, 1997).

However, the EADS is not only a means of communicating to forecast users. It also impacts on the forecasting process itself, as all forecasters are working to an agreed, common, and at least nominally binding, definition of avalanche hazard.

Forecast validation and evaluation is not only a problem in avalanche forecasting, but more generally in forecasting. Murphy (1993), in his classic paper on the nature of a good (weather) forecast, discussed three key elements which he termed *consistency*, *quality* and *value*. Consistency in Murphy's model essentially captures the degree of agreement between a forecaster's understanding of a situation and the forecast they then communicate to the public. Quality captures the degree of agreement between a forecast and the events which occur, and value the benefits or costs incurred by a user as a result of a forecast.

In avalanche forecasting, two key problems come to the fore. Firstly, the target variable is essentially categorical, since although the EADS is an ordinal scale, a real evaluation of a forecast would compare the forecast danger level, qualitatively defined in the EADS, with the prevailing avalanche situation. Secondly, since the target variable captures a state which may or may not lead to an (avalanche) event, verification of forecast quality is only possible in some circumstances and for some aspects of the EADS, for example:

- At higher danger levels, the occurrence of natural avalanches can sometimes be used to verify the danger level (e.g. Elder and Armstrong, 1987; Giraud et al., 1987; Schweizer et al., 2018).
- At lower danger levels, the occurrence of avalanches triggered by recreationists or the observation of signs of instability requires users being present.
- Since the absence of avalanche activity is not alone an indicator of stability, verifying associated danger levels is only possible through digging multiple snow profiles and performing stability tests (Schweizer et al., 2003).

Thus, avalanche danger cannot be fully measured or validated (Föhn and Schweizer, 1995). This in turn means that, at least at the level of the EADS, it is conceptually difficult to directly measure forecast quality. However, Murphy's notion of considering goodness of forecasts in terms of not only their quality, but also consistency and value, suggests a possible way forward.

Although Murphy defines consistency with respect to an individual forecaster, we believe that the concept can be extended to forecast centers, in terms of the degree to which individual forecasters using potentially different evidence reach the same judgment (LaChapelle, 1980), and across forecast centers, in terms of the uniformity of the forecast issued by different forecast centers in neighboring regions. This reading of consistency is, we believe, both true to Murphy's notion (how reliably does a forecast correspond with a forecaster's best judgment) and broader notions of consistency stemming from work on data quality and information science (Ballou and Pazer, 2003; Bovee et al., 2003).

Inconsistencies in the use of the danger levels between neighboring warning regions and forecast centers may be a potential source of misinterpretations to users traveling from one region to another, unless these differences are only due to avalanche conditions. The main goal of this study is therefore to investigate if such spatial inconsistencies and biases exist. We do so by quantifying bias between neighboring forecast centers and warning regions in time and space. While we do not expect spatial homogeneity, a stronger bias and a lower agreement rate in neighboring warning regions in different forecast centers, compared to within forecast domains, may indicate such inconsistencies. To do so, we first describe the operational constraints under which avalanche forecasts are produced and communicated. Then, we present methods appropriate to explore spatial consistency and bias in the use of EADS given the operational constraints described above. We address the following three research questions:

1. Does bias between forecast centers exist?
2. Can operational constraints (such as the size of the warning regions) or the elevation of warning regions explain these differences?
3. What implications do the biases identified have for users of avalanche forecasts?

## 2 Background and definitions

In the following, we introduce the most important standards, concepts and definitions used in avalanche forecast products in the European Alps. We describe the situation during the winters 2011/2012 until 2014/2015, as these are the years we explore quantitatively in this study.

### 2.1 Avalanche warning services and forecast centers

Avalanche warning services (AWS) are national, regional or provincial agencies in charge of providing publicly available forecasts of avalanche hazard (EAWS, 2017c). AWS also have voting rights at the General Assembly of the European AWS (EAWS). An AWS may either be a service with a single forecast center (e.g. the national service in Switzerland or the regional AWS of the federal states in Austria) or with several forecast centers in different locations (e.g. the AWS Météo-France in France with four forecast centers in the Alps or the two AWS in Italy (Associazione Interregionale Neve e Valanghe (AINEVA) and Meteomont Carabinieri) with their provincial and regional centers.

Generally, and with the exception of Italy, a single forecast covering a (number of) warning region(s) is issued by the respective forecast center (Tab. 2, Fig. 1). In the case of Italy, forecast centers belonging to AINEVA and Meteomont Carabinieri independently provide forecasts covering the same Alpine regions, while in Livigno (LIV in Fig. 1) a regional forecast is also issued by the municipality. Even though the forecast products provided by the individual forecast centers may differ in their structure, we assume they adhere to the principles defined by the European Avalanche Warning Services (EAWS, 2017c).

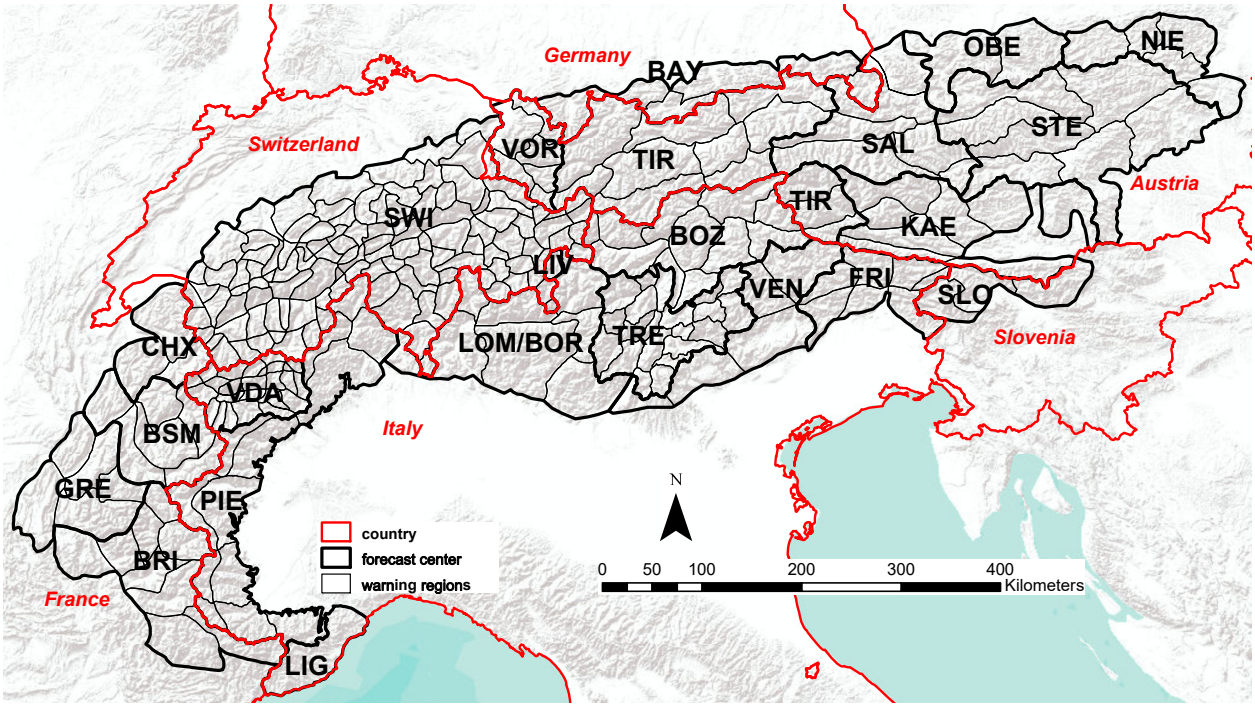


**Figure 1.** Map showing the relief of the European Alps (gray shaded background) with the outlines of the individual forecast centers (bold black polygons, three-letter abbreviations) and the warning regions, the smallest geographically defined regions, used in the respective avalanche forecasts (black polygons). The borders of the Alpine countries are marked red. In the Italian Alps, where two avalanche warning services provide forecasts (Associazione Interregionale Neve e Valanghe (AINEVA) and Meteomont Carabinieri), the warning regions generally follow AINEVA. An exception is LIG (avalanche warning service Meteomont Carabinieri). The forecast domains of LOM (AINEVA) and BOR (Meteomont Carabinieri) are identical, however, the three warning regions for BOR are not shown on the map. The forecast domain LIV is superposed onto parts of LOM/BOR (map source: ESRI, 2017). Note that the map captures the situation and partitioning during the period under study.

**Table 1.** European avalanche danger scale (EAWS, 2018).

| Danger level | Snowpack stability | Likelihood of triggering |
|---|---|---|
| 5-Very High | The snowpack is poorly bonded and largely unstable in general. | Numerous very large and often extremely large natural avalanches can be expected, even in moderately steep terrain*. |
| 4-High | The snowpack is poorly bonded on most steep slopes*. | Triggering is likely even by low additional loads** on many steep slopes*. In some cases, numerous large and often very large natural avalanches can be expected. |
| 3-Considerable | The snowpack is moderately to poorly bonded on many steep slopes*. | Triggering is possible even from low additional loads** particularly on the indicated steep slopes*. In certain situations some large, in isolated cases very large natural avalanches are possible. |
| 2-Moderate | The snowpack is only moderately well bonded on some steep slopes*; otherwise well bonded in general. | Triggering is possible primarily from high additional loads**, particularly on the indicated steep slopes*. Very large natural avalanches are unlikely. |
| 1-Low | The snowpack is well bonded and stable in general. | Triggering is generally possible only from high additional loads** in isolated areas of very steep, extreme terrain**. Only small and medium-sized natural avalanches are possible. |

* The avalanche-prone locations are described in greater detail in the avalanche bulletin (altitude, slope aspect, type of terrain): moderately steep terrain: slopes shallower than about 30 degrees; steep slopes: slopes steeper than about 30 degrees
very steep, extreme terrain: particularly adverse terrain related to slope angle (more than about 40 degrees), terrain profile, proximity to ridge, smoothness of underlying ground surface
** Additional loads:
low: individual skier / snowboarder, riding softly, not falling; snowshoer; group with good spacing (minimum 10m) keeping distances
high: two or more skiers / snowboarders etc. without good spacing (or without intervals); snowmachine; explosives
natural: without human influence

**Table 2.** Overview of the forecast centers considered in this study. Italian forecast centers refer to AINEVA, except those indicated with subscript MC for Meteomont Carabinieri. Forecast centers and warning regions outside the Alps are not shown. Three-letter abbreviations indicate forecast centers. For countries, we use English names, for forecast centers the names in their original language.

| country | forecast center | abbreviation | surface area* in $\mathrm{km}^2$ | number of warning regions | size** median (min-max) in $\mathrm{km}^2$ | max. elevation*** min-max in $\mathrm{m.a.s.l.}$ |
|---|---|---|---|---|---|---|
| Austria | Kärnten | KAE | 7700 | 8 | 1060 (520 - 1300) | 2110 - 3740 |
| | Niederösterreich | NIE | 3700 | 5 | 730 (500 - 1030) | 1390 - 2060 |
| | Oberösterreich | OBE | 3400 | 2 | 1720 (1530 - 1910) | 2360 - 2860 |
| | Salzburg | SAL | 6800 | 6 | 1090 (360 - 1970) | 2010 - 3570 |
| | Steiermark | STE | 12500 | 7 | 2030 (1250 - 2290) | 1770 - 2800 |
| | Tirol | TIR | 12600 | 12 | 980 (380 - 1920) | 2460 - 3730 |
| | Vorarlberg | VOR | 2600 | 6 | 390 (180 - 880) | 2080 - 3200 |
| Switzerland | Schweiz | SWI | 26300 | 117 | 180 (40 - 660) | 1640 - 4550 |
| Germany | Bayern | BAY | 4300 | 6 | 660 (450 - 1190) | 1870 - 2940 |
| France | Bourg-St-Maurice | BSM | 5100 | 6 | 810 (630 - 1220) | 2160 - 3810 |
| | Briançon | BRI | 8000 | 9 | 840 (450 - 1590) | 2760 - 4020 |
| | Chamonix | CHX | 3000 | 3 | 1070 (580 - 1380) | 2700 - 4780 |
| | Grenoble | GRE | 5300 | 5 | 990 (560 - 1440) | 2070 - 3950 |
| Italy | Bozen-Südtirol / Bolzano-Alto Adige | BOZ | 7400 | 11 | 650 (180 - 1110) | 2590 - 3860 |
| | Friuli Venezia Giulia | FRI | 3700 | 7 | 560 (160 - 690) | 1880 - 2740 |
| | Liguria and Toscana$^{MC}$ | LIG | 2100 | 1 | 2060 | 2140 |
| | Livigno | LIV | 200 | 1 | 210 | 3210 |
| | Lombardia | LOM | 9700 | 7 | 1330 (510 - 2820) | 2230 - 3940 |
| | Lombardia$^{MC}$ | BOR | 9700 | 3 | 3120 (1900-4630) | 2850 - 3940 |
| | Piemonte | PIE | 10300 | 13 | 820 (270 - 1630) | 2580 - 4530 |
| | Trentino | TRE | 6200 | 21 | 290 (120 - 540) | 2060 - 3620 |
| | Valle d'Aosta | VDA | 3300 | 26 | 130 (25 - 280) | 2620 - 4780 |
| | Veneto | VEN | 5500 | 5 | 1100 (460 - 1640) | 2180 - 3250 |

* rounded to nearest 100 $\mathrm{km}^2$, ** rounded to nearest 10 $\mathrm{km}^2$, *** rounded to nearest 10 $\mathrm{m}$
The size, as shown here and in Figure 3a, was calculated using the R-package *raster* (Hijmans, 2016).
The range of the maximum elevations describes the range of the highest elevation calculated using a digital elevation model with 90x90 $\mathrm{m}$ cell resolution (Jarvis et al., 2008; SRTM, 2017) per warning region and forecast center (Figure 3b).

### 2.2 Avalanche forecasts

Avalanche forecasts are the primary means for avalanche warning services to provide publicly available information about current and forecast snow and avalanche conditions in their territory. They may take the form of a single advisory, describing the current situation, or an advisory and forecast for one or more days. Typically, avalanche forecasts contain the following information, ranked according to importance (*information pyramid*; EAWS, 2017b):

- avalanche danger level according to the EADS (Table 1)
- most exposed terrain - defining the terrain where the danger is particularly significant (see section 2.5).
- typical avalanche problems - describing typical situations encountered in avalanche terrain (EAWS, 2017d)
- hazard description - a text description providing information concerning the avalanche situation
- information concerning snowpack and weather

In this study, we exclusively explore the forecast regional avalanche danger level. However, we also describe how the danger level is communicated in relation to the most exposed terrain (by elevation) and to its temporal evolution during the day, as this differs between forecast centers and could influence the results.

### 2.3 Temporal validity and publication frequency

The issuing time, temporal validity and publication frequency of the forecasts varies between forecast centers. For the explored four winters, these can roughly be summarized in five groups (the «normal» cases are described, exceptions exist; see also Fig. 2):

1. Bulletins are published daily in the morning (generally around 07:30 CET[2]) and are normally valid for the day of publication (typical for bulletins in Austria, Germany and Livigno (LIV/Italy).
2. Forecasts are published daily in the afternoon (16:00 CET) and are valid until the following day (France).
3. During the main winter season (often from early December until after Easter), forecasts are published twice daily. The main forecast, published at 17:00 CET valid until 17:00 CET the following day, is replaced by an update the following morning at 08:00 CET (Switzerland).
4. Bulletins are published several times a week (at least on Monday, Wednesday, Friday). Bulletins are issued between 11:00 and 17:00 CET and describe the avalanche conditions on the day of publication, the following day and the day after (typically forecast centers belonging to AWS AINEVA). In more recent years, publication frequency increased towards daily.
5. Bulletins are published at 14:00 CET, describing the current situation and the forecast for the next day(s). Forecasts are published daily, except on public holidays (AWS Meteomont Carabinieri).

Most of the forecast centers can update their forecast product when conditions change significantly.

### 2.4 Warning regions

Warning regions are geographically clearly specified areas permitting the forecast user to know exactly which region is covered by the forecast. They may be delineated by administrative boundaries (e.g. between countries, federal states, or regions and provinces), describe climatologically (e.g. in France; Pahaut and Bolognesi (2003)), hydrologically or meteorologically homogeneous regions, or may be based on orographic divisions (e.g. Italy; Marazzi, 2005), or a combination of these (e.g. Valle d'Aosta (VDA); Burelli et al., 2012). Generally, warning regions are larger than the minimal spatial resolution of a regionally forecast avalanche danger level, and are therefore recommended to have a size of about 100 km$^2$ or larger (EAWS, 2017c).

The median size of the warning regions is 350 $\mathrm{km}^2$ with considerable variations (Fig. 3a, Tab. 2). The 25% of the smallest warning regions (size $<$ 160 $\mathrm{km}^2$, all located in Switzerland (SWI), Trentino (TRE) and Valle d'Aosta (VDA)) are almost ten times smaller than the 10% of the largest regions (size $>$ 1310 $\mathrm{km}^2$). Particularly large spatial units are used by the forecast centers covering the region of Lombardia (BOR) and the Ligurian Alps (LIG, both AWS Meteomont Carabinieri, Italy) and in Oberösterreich (OBE, Austria; size $>$ 1900 $\mathrm{km}^2$, Table 2). The size of the warning regions depends on the approach used by an AWS to define the warning regions and to externally communicate avalanche danger. In its simplest case (see variations introduced in next section), a single danger level is either explicitly communicated for each warning region (e.g. in Austria, France, Germany, often in Italy) or may be communicated for an aggregation of warning regions (Switzerland (SWI), Trentino (TRE) and Valle d'Aosta (VDA)).

[2] all times indicated may refer to either CET or CEST

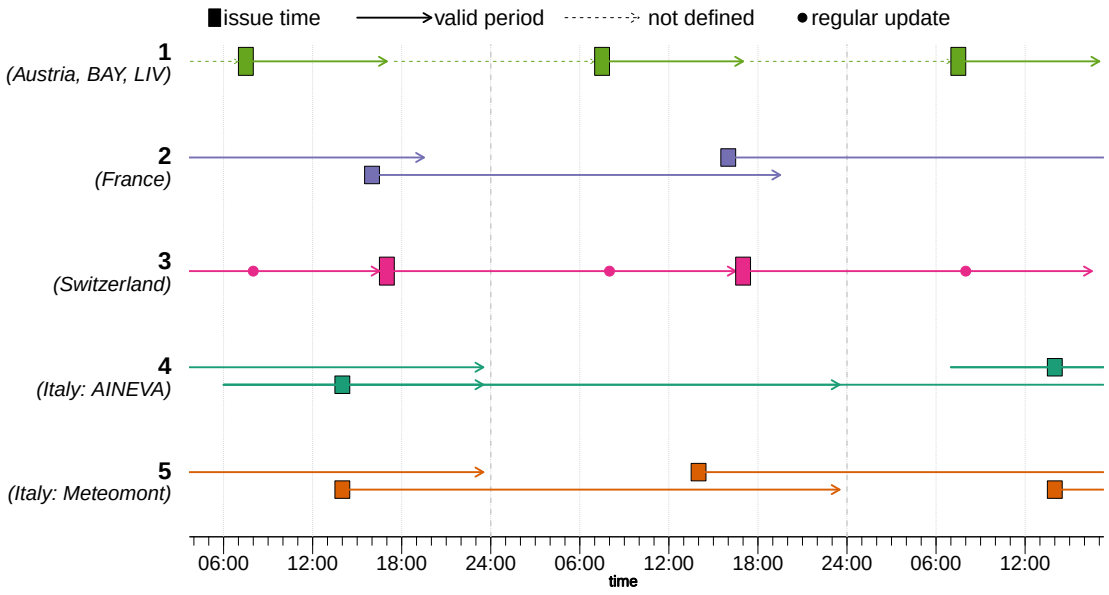


**Figure 2.** Schematic summary of the different bulletin publication frequencies, issuing times and periods of validity. In special circumstances, updates during the morning were possible in most forecast centers. Particularly for Italy (AINEVA), it is of note that the exact publication times, valid periods and publication frequencies may differ between forecast centers, but changes may also have been introduced from one season to the next. Forecast centers are labeled according to Table 2.

### 2.5 Concepts to communicate temporal changes and elevational gradients in danger level

The communication of the most exposed elevations and slopes, and expected temporal changes are important information provided in avalanche forecasts.

#### 2.5.1 Temporal differences in danger rating within forecast period

All forecast centers communicate significant changes (increasing or decreasing danger level) during the valid period of a forecast. In most cases, this is done graphically using either icons or two maps, and only rarely using text. In cases, when two danger levels are indicated, the first time-step often refers to the avalanche danger in the morning, the second time-step indicates a significant change during the day. Changing danger ratings may refer to either changes in dry- or wet-snow avalanche hazard, or from dry- to wet-snow (or vice versa). However, exceptions to these generalizations exist: In France, but occasionally in forecasts of other forecast centers too, the two time-steps may refer to either day and night, morning or afternoon, or before and after a snowfall. Switzerland is the only warning service where an increase in danger rating for wet-snow situations (typically in spring conditions) is presented using a map product if the wet-snow rating is higher than the dry-snow rating in the morning, but an increase in dry-snow avalanche hazard during the day is exclusively conveyed in text form within the danger description.

#### 2.5.2 Elevational differences in danger rating

All forecast centers provide information concerning the most exposed elevations, often in graphical form using icons. The elevational threshold indicated in the bulletin may relate to a difference in danger rating (for instance higher above a certain elevation), or differences in the avalanche problem and the most likely type of avalanche expected (e.g. wet-snow avalanches below and dry-snow avalanches above the indicated elevation), or a combination thereof.

The forecast centers use three different ways to communicate elevational differences in the danger rating. In Switzerland and Italy, the danger rating refers to the most exposed elevations, with no indication of the (lower) danger rating in other elevations. In France, Germany and some regions in Austria, two separate danger ratings are often provided: one above a certain elevation level, and one below, while the forecast center Livigno (LIV) in northern Italy assigns a danger rating to the three elevation bands *below tree line*, *tree line* and *alpine* (as done in North American avalanche forecasts).

## 3 Data

We approached all the warning services in the Alps concerning the forecast danger level for each warning

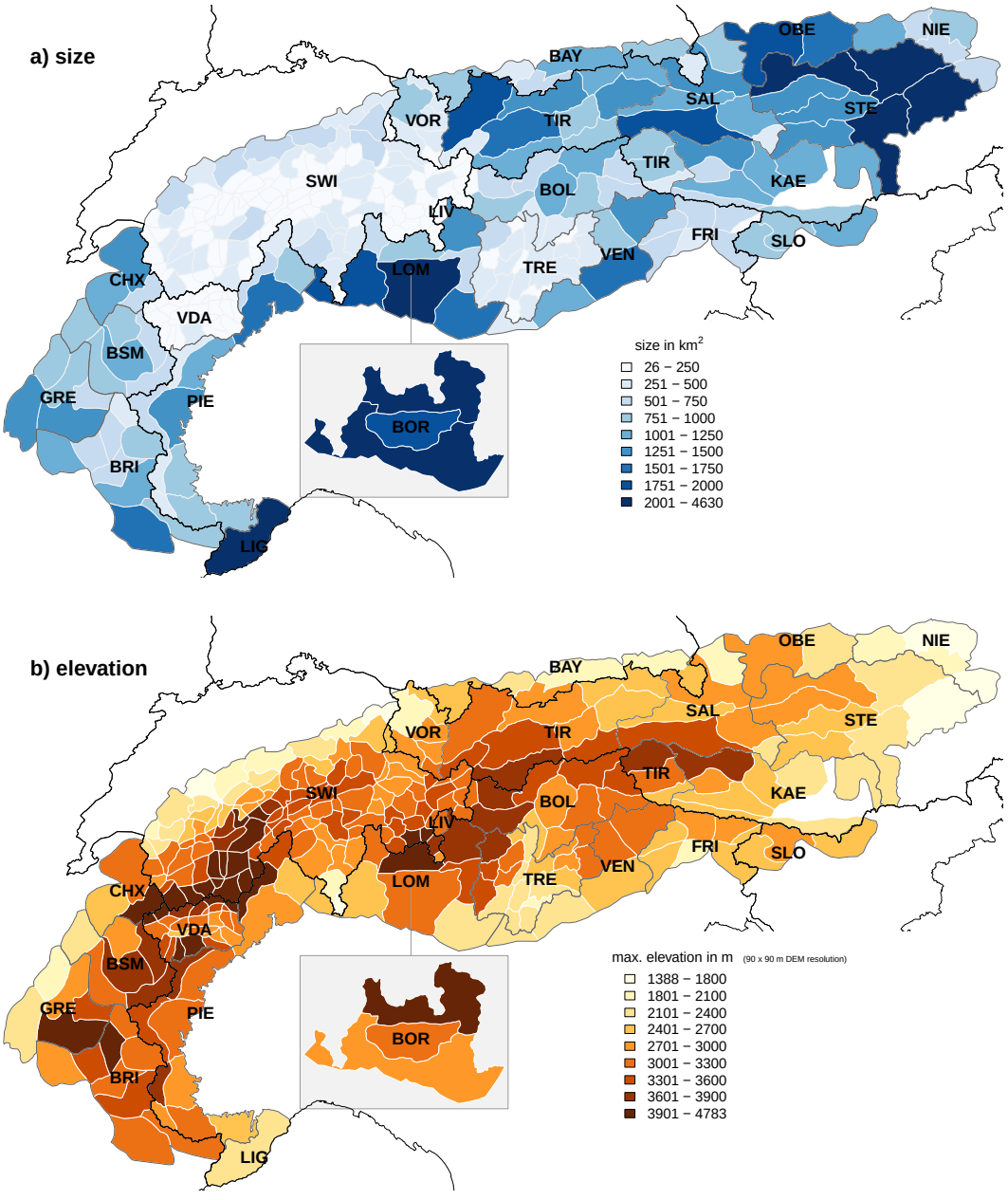


**Figure 3.** Map showing the European Alps with the individual warning regions (white polygon outlines) and (a) their size (color shading of polygon) and (b) their maximum elevation (color shading of polygon). Additionally, national (black lines) and forecast center boundaries (grey polygon outlines) are shown. To visualize the (at least partially) overlapping forecast regions in the Italian region of Lombardia, LIV is superposed onto parts of LOM, while BOR is placed as inset to the south of LOM. Forecast centers are labeled according to Table 2.

region and day for the four years from 2011/12 to 2014/15 and received data from 23 of the 30 forecast centers.

### 3.1 Avalanche danger level data

In most cases, data were provided directly from the warning services or forecast centers. Exceptions were

- Kärnten (KAE, Austria) - Data were extracted from the annual reports ÖLWD (2012)-ÖLWD (2015)
- Bayern (BAY, Germany) - Data were collected from the web archive of the Bavarian warning service
- AINEVA forecast centers Friuli Venezia Giulia (FRI), Lombardia (LOM), Veneto (VEN) in Italy - Data were

provided by M. Valt/VEN (extracted from the central AINEVA database).

The most relevant information concerning differences in raw data analyzed are displayed in Tab. 3. The danger level was generally valid for the day of publication (d+0, in Austria, Germany, Livigno (LIV), scenario 1 in Sect. 2.3), represented essentially a one-day forecast (d+1) in France and Switzerland (although the valid period started already on the afternoon of publication, scenario 2 and 3 in Sect. 2.3), but was a mix of current day assessments, and forecasts with one or two days (d+2) lead-time in Italy. In Italy (AWS AINEVA), the most recently published valid danger level was used (e.g. an afternoon update, valid for the current day (d+0) replaced a forecast with a lead time of two days (d+2))). Furthermore, publication frequency increased during the explored time period in some of the AINEVA forecast centers (i.e. in Piemonte (PIE) to weekdays or in Bozen-Südtirol/Bolzano-Alto Adige (BOZ) additionally on Saturdays). Similarly, the validity of the bulletin on the issuing day changed in some Italian forecast centers from a current day assessment to a one-day forecast (i.e. BOZ changed in 2014 from d+0 to d+1), or vice versa (AWS Meteomont Carabinieri: Lombardia (BOR) and Liguria e Toscana (LIG) changed in 2014 from d+1 to d+0).

Temporal differences in danger level within the forecast period were available for all forecasts, except those by BOR and LIG (Italy) and KAE (Austria). In both cases, only the highest danger level per day was available. The data extracted from the AINEVA database (forecast centers Friuli Venezia Giulia (FRI), Lombardia (LOM), Veneto (VEN)) indicated not only the danger level, but also whether the danger rating increased, stayed the same or decreased.

In France spatial variations in the danger level within the same warning region ($D_{spatial}$) were sometimes indicated (e.g. in a bulletin this could read «2-Moderate in the West, 3-Considerable in the East» of a region).

# 4 Methods

The quantitative part of this study is twofold: first, we make pairwise comparisons of neighboring warning regions, and second, we visualize and detect patterns at larger scales than individual warning regions.

## 4.1 Topological neighbors

We defined warning regions *i* and *j* as topological neighbors, whenever they shared more than one point of their polygon boundary with each other (rook mode, Dale and Fortin (2014); R-package *spdep* Bivand et al. (2013); Bivand and Piras (2015)). For this purpose, the shapes of the warning regions had to be slightly adjusted so that the coordinates of joint borders matched. This also reflects challenges of working across borders, with different map projections and simplified outlines of warning regions. For the particular case of the three forecast centers in Lombardia (BOR, LIV and LOM), we defined them as neighbors if they either shared a common polygon boundary or at least partially the same territory.

## 4.2 Avalanche danger level statistics

We refer to danger levels D either using their integer value (e.g. D=1 for *1-Low*) or by integer value and signal word combination *1-Low*. Similarly to previous studies (e.g. Jamieson et al., 2008; Techel and Schweizer, 2017), we use the integer value of danger levels to calculate proportions and differences.

### 4.2.1 Data preparation

We explored the forecast danger levels at the spatial scale of the individual, geographically clearly delineated warning regions. The following cases were treated separately:

**Austria, Germany, France**: occasional updates during the morning

In special circumstances, bulletins were updated during the day and the danger level adjusted. These cases were rare (for instance in Bayern (BAY) and Tirol (TIR) twice during the explored four winters). These updates were not considered in the analysis. The data provided by France, where morning updates are also possible until 10:00 CET, already included such updates.

**France:** spatial gradients within same warning region

In France, forecasters sometimes communicated two danger ratings for the same warning region expressing a spatial gradient. These cases were rare (0.4% of warning regions and days; Bourg-St-Maurice (BSM) 1%, Briançon (BRI) 0.3%, Chamonix (CHX) 0.1%, Grenoble (GRE) 0%). For these forecasts, we randomly picked one of the two danger levels. The remainder of the forecasts expressed no spatial gradients.

**Switzerland (SWI):** evening forecast; danger ratings communicated in text form only

We used the forecast issued at 17:00 CET, rather than the updated forecast the next morning (08:00 CET) as, until the winter 2012/13, the daily morning update was issued only for parts of the Swiss Alps. Furthermore, we only analyzed the danger ratings published on the map product, and not those only described in the forecast text (Section 2.5).

**Italy (AINEVA forecast centers Friuli Venezia Giulia (FRI), Lombardia (LOM), Veneto (VEN))**: forecast danger level changed during valid bulletin period

Data extracted from the AINEVA database provided the danger level valid in the morning, and whether the danger

**Table 3.** Overview of the data used in this study. Forecast centers are summarized according to data source, format and content. $D_{t1}$ and $D_{t2}$ - danger level time step 1 and 2, respectively; $D_e$ - concept of elevational danger ratings; $D_{spatial}$ - more than one rating per warning region referring to spatial differences. Danger levels may refer to the day of publication (day+0), the following day (day+1) or the day after (day+2). Forecast centers are labeled according to Table 2.

| country | forecast center | $D_{t1}$ | $D_{t2}$ | $D_e$ | $D_{spatial}$ | day+0 | day+1 | day+2 | source |
|---|---|---|---|---|---|---|---|---|---|
| Austria | KAE | no | yes | 2 | no | 100% | – | – | ÖLWD |
| | NIE, OBE, SAL, STE, TIR, VOR | yes | yes | 2 | no | 100% | – | – | directly |
| Switzerland | SWI | yes | yes | 1 | no | – | 100% | – | directly |
| Germany | BAY | yes | yes | 2 | no | 100% | – | – | website |
| France | BSM, BRI, CHX, GRE | yes | yes | 2 | yes | – | 100% | – | directly |
| Italy | BOZ, PIE, TRE, VDA | yes | yes | 1 | no | 42% | 41% | 16% | directly |
| | FRI, LOM, VEN | yes | (yes)* | 1 | no | – | – | – | AINEVA |
| | BOR, LIG | no | yes | 1 | no | 48% | 49% | 3% | directly |
| | LIV | yes | yes | 3 | no | 100% | – | – | directly |

* (yes): AINEVA database provided information whether danger level changed, but not to which danger level
$D_e$: concept of assigning 1, 2 or 3 danger ratings (Sect. 2.5)
data source: ÖLWD - from Austrian winter reports ÖLWD (2012) - ÖLWD (2015), directly - directly from respective forecast center, website - from website of Bavarian avalanche warning service, AINEVA - extracted from central AINEVA database (M. Valt (VEN))

level changed during the day (increase, no change, decrease), but not which danger level was forecast following the change. To supplement this information, we utilized the distributions of the four AINEVA forecast centers, which consistently provided the second danger rating (Bozen-Südtirol/Bolzano-Alto Adige (BOZ), Piemonte (PIE), Trentino (TRE), Valle d'Aosta (VDA)). In these forecasts, changing danger level was by one level in 85% of cases, and by two levels in 15% of cases. For the bulletins in FRI, LOM and VEN we assumed a one-level difference for days with changing conditions, and hence a somewhat more conservative value than in the other Italian bulletins.

**Standardizing the length of the forecasting period during the season**

The length of the main forecasting season is considered as being between 14 December and 16 April. During this time, and with the exception of the 2014/15 winter (28 Dec - 16 April), there was a danger rating in at least 95% of the warning regions in the Alps (477 days, 4 winters).

#### 4.2.2 Danger ratings $D_{max}$ and $D_{morning}$

We created two subsets of data ($D_{max}$ and $D_{morning}$), to accommodate the different ways avalanche danger ratings are communicated in forecasts and stored in databases, and to ascertain that no bias was introduced by these differences.

We defined $D_{max}$ as the highest danger rating valid during a forecast period, regardless whether this was the only rating provided, whether this was for a first or second time-step, or whether it corresponded to a difference in danger level by elevation. It is of note, that $D_{max}$ is sometimes only valid for part of the day or part of the elevation range.

In contrast, $D_{morning}$ refers to the maximum danger rating for the first of the two time steps, which in many cases would be considered valid for the morning. Here, it is of note that exact time when a change occurs is never provided in the published forecasts, and only categorically described within the danger description. This was calculated for all forecast centers, except Lombardia (BOR), Liguria (LIG) and Kärnten (KAE), where this information was not available.

#### 4.2.3 Summary statistics

**Warning region-specific summary statistics**

For each warning region, we calculated the proportion of forecasts issuing a specific danger level (i.e. forecasts with danger level $D = 4$). Furthermore, for each warning region we calculated the surface area, which we refer to as the *size* of a warning region, using the R-package *raster* (Hijmans, 2016) and the maximum elevation (ArcGIS software). The latter is based on a 90x90 $\mathrm{m}$ digital elevation model (ESRI, 2017).

**Pairwise comparison of immediately neighboring warning regions**

We compare the forecast danger level in two neighboring warning regions $i$ and $j$ by calculating the difference in the forecast danger level $\Delta D$ for each day $\Delta D = D_i - D_j$ for all days with $D_i \geq 1$ and $D_j \geq 1$, where D may refer to $D_{max}$ or $D_{morning}$.

The proportion of days when the forecast danger levels

agreed $P_{agree}$ is then

$$P_{agree} = P(\Delta D = 0) = \frac{N(\Delta D = 0)}{N(\Delta D)} \quad (1)$$

$P_{agree}$ may be interpreted as an indicator of spatial correlation or measure of spatial continuity in avalanche conditions.

For neighboring warning regions $i$ and $j$, we calculated a bias ratio $B_{ij}$ similar to Wilks (2011, p. 310):

$$B_{ij} = \frac{N(\Delta D = 0) + N(\Delta D^{+})}{N(\Delta D = 0) + N(\Delta D^{-})} \quad (2)$$

where $N(\Delta D^{+})$ is the number of days with the $D_i \geq D_j$ and $N(\Delta D^{-})$ the number of days with $D_i \leq D_j$. $B_{ij} > 1$ indicates region $i$ having more frequently higher danger levels than region $j$, $B_{ij} = 1$ indicates a perfectly balanced distribution, and $B_{ij} < 1$ a skew towards more often higher danger levels in region $j$ compared to $i$. We tested whether the bias $B_{ij}$ was significantly unbalanced, by comparing the observed distribution of the two outcomes ($N(\Delta D^{+})$, $N(\Delta D^{-})$) to a random distribution using the binomial test (R: *binom.test*, R Core Team (2017)). The resulting p-value depends on the deviation of $B_{ij}$ from 1, and on the number of days $N(\Delta D \neq 0)$. In general, bias values $B_{ij} < 0.95$ or $B_{ij} > 1.05$ were statistically significant ($p < 0.05$).

The distance between warning regions refers to the distance between the center points of the respective warning regions.

**Sensitivity and correlation**

We tested whether removing subsets of the data (for instance individual years), or using $D_{morning}$ compared to $D_{max}$ influenced the rank order of the warning regions using the Spearman rank order correlation coefficient $\rho$. Similarly, we used $\rho$ to explore whether the frequency a specific danger level was issued correlated with differences in the size ($\Delta_{size}$) or in the maximum elevation ($\Delta_{elevation}$) of two warning regions $i$ and $j$.

We compared populations using the Wilcoxon rank-sum test (Wilks, 2011, p. 159-163). We consider $p \leq 0.05$ as significant.

# 5 Results

## 5.1 Forecast danger levels

Fig. 4 summarizes the distribution of issued danger levels across the Alps during the four years (477 forecast days, 281 warning regions). Danger levels *2-Moderate* and *3-Considerable* are forecast about 80% of the time, regardless whether we consider the forecast danger level valid in the first time-step, often corresponding to the situation in the morning ($D_{morning}$; Fig. 4a), or the highest danger level issued ($D_{max}$; Fig. 4b). Particularly in spring situations, when avalanche hazard often increases with day-time warming, the afternoon rating is higher than the morning one; hence these two distributions differ significantly ($p < 0.01$). However, as often the results obtained using $D_{max}$ and $D_{morning}$ were very similar, in the following we only present results if these differed significantly.

In order to address research questions 1 and 2, we explore agreement and bias (Sect. 5.2), the proportion of forecasts at the upper and lower end of the EADS (Sections 5.3 and 5.4), and the proportion of changing danger ratings during the day (Sect. 5.5). Additionally, we explore the influence of the size of the warning regions on the spatial variability in danger ratings and on the proportion of forecasts with danger levels *4-High* and *5-Very High* (Section 5.6). Finally, we present two case studies to illustrate different aspects of these results in practical situations (Section 5.7).

## 5.2 Comparing immediately neighboring warning regions: agreement and bias

The forecast danger level agreed in 83% of the cases (median $P_{agree}$) between two neighboring warning regions. $P_{agree}$ was significantly higher when comparing warning regions within forecast center boundaries (91%, interquartile range IQR 83 - 96%) compared to those across forecast center boundaries (63%, IQR 58 - 70%, $p < 0.001$), or across national borders (62%, IQR 58 - 66%, $p < 0.001$). The latter values were not significantly different. Exploring the agreement rate graphically on a map by emphasizing borders with $P_{agree} \leq 80\%$ essentially captures almost all forecast center boundaries and comparably few boundaries within forecast center domains (Fig. 5). This result is confirmed when using only a subset of the warning region pairs, with $\Delta_{elevation} < 250\ \mathrm{m}$ and the size of the larger region being less than 1.5 times the size of the smaller region (Fig. 6). For this subset, the median agreement $P_{agree}$ is about 30% lower across forecast center boundaries, than within those ($P_{agree}$(same forecast center) = 93%, $P_{agree}$(different forecast center) = 63%, $p < 0.001$, Fig. 6). Even when removing the data of the forecast centers in Switzerland (SWI), Trentino (TRE) and Valle d'Aosta (VDA), with median $P_{agree}$ values of 95%, the difference remains highly significant ($P_{agree}$(within forecast center domain) = 87%, $P_{agree}$(across forecast center domains) = 63%, $p < 0.001$).

Similar results are noted for the special case of the three forecast centers in the Italian region of Lombardia (BOR, LIV, LOM). For these partially overlapping warning regions $P_{agree}$ was 63%, and thus similar to $P_{agree}$ across national

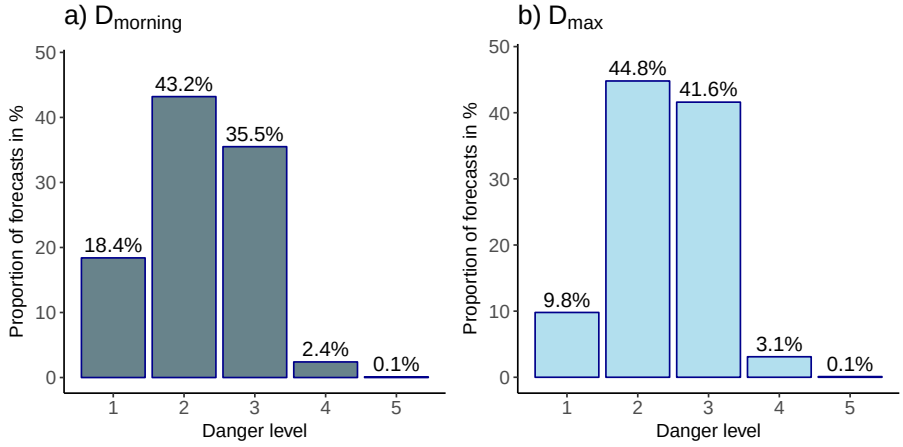


**Figure 4.** Distribution of forecast danger levels, for a) $D_{morning}$ (danger level valid during first time-step) and b) $D_{max}$ (highest danger level). Mean values are shown for all the warning regions in the Alps taken together.

borders or forecast centers neighboring each other.

Within the boundaries of forecast centers, there was a weak, but significant correlation between $P_{agree}$ and differences in the elevation of two neighboring regions ($\rho$ = -0.36, p $<$ 0.001), with larger differences in elevation corresponding to a lower agreement rate. There was also a weak correlation between $P_{agree}$ and differences in the size of the warning regions ($\rho$ = -0.24, p $<$ 0.001), where agreement increases as the size difference between warning regions decreases.

Within forecast center domains, the bias ratio $B_{ij}$ correlated weakly with differences in the size ($\rho$ = -0.37, p < 0.01) and elevation ($\rho$ = -0.21, p < 0.01), indicating that generally the forecast danger level increased with elevation, but also with the size of the warning region. For the warning regions pairs shown in Fig. 6, a significant bias existed in 76% of the pairs across forecast center boundaries, compared to 51% within those boundaries.

Compared to warning regions in neighboring forecast centers, the forecast centers Niederösterreich (NIE), Switzerland (SWI) and Bayern (BAY) had the lowest median bias ratios ($B_{ij} \leq 0.84$), indicating that lower danger levels were used more frequently. This is in contrast to Lombardia (LOM), Briançon (BRI) and Salzburg (SAL) with median bias ratios $B_{ij} \geq 1.19$. For days and regions where danger levels differed, this corresponded to $D_{max}$ being lower on more than two thirds of the pairwise-comparisons for Niederösterreich (NIE), Switzerland (SWI) and Bayern (BAY), and similarly for Lombardia (LOM), Briançon (BRI) and Salzburg (SAL) with more than 60% of forecasts with $\Delta D \neq 0$ being higher.

### 5.3 Very critical avalanche conditions D $\geq$ 4

Danger level *5-Very High* was rarely forecast (less than 0.1% of days and regions, mostly during 2013/2014 in the southern part of the Alps; Fig. 4). Therefore, we explore forecasts with a very critical avalanche situation (D=4) or a disaster situation (D = 5) combined. For a specific warning region, the proportion of forecasts with very critical conditions is

$$P_{\text{v.crit}} = \frac{N(D \geq 4)}{N} \tag{3}$$

where N is the number of forecasts.

Forecasts with forecast danger levels *4-High* or *5-Very High* were generally rare (median 2.5%, IQR: 1.1 - 4%, Fig. 7), but were considerably more frequently forecast in the warning regions belonging to the four forecast centers in France (Briançon (BRI), Bourg-St-Maurice (BSM), Chamonix (CHX), Grenoble (GRE)) and the Italian forecast centers Piemonte (PIE) and Lombardia (LOM). Visually exploring spatial patterns (Fig. 7a) shows several forecast center borders which coincide with large gradients in $P_{v.crit}$ values. These differences are most obvious when comparing Switzerland (SWI) with its neighbors Chamonix (CHX), Piemonte (PIE), Lombardia (LOM) and Tirol (TIR), where two (or more) classes difference often occur. In contrast, and with some exceptions, comparably similar values can be noted in many of the forecast centers in Austria, Germany, Switzerland and the Italian provinces and regions of Valle d'Aosta (VDA), Bozen-Südtirol/Bolzano-Alto Adige (BOZ) and Trentino (TRE). Variations are also confirmed, when considering only warning regions with a maximum elevation greater than 2500 $\mathrm{m}$ (N=222). Median values for warning regions in Bozen-Südtirol/Bolzano-Alto Adige (BOZ), Switzerland (SWI), Vorarlberg (VOR), Valle d'Aosta (VDA) and Salzburg (SAL) (1.6%-2.3%) are significantly

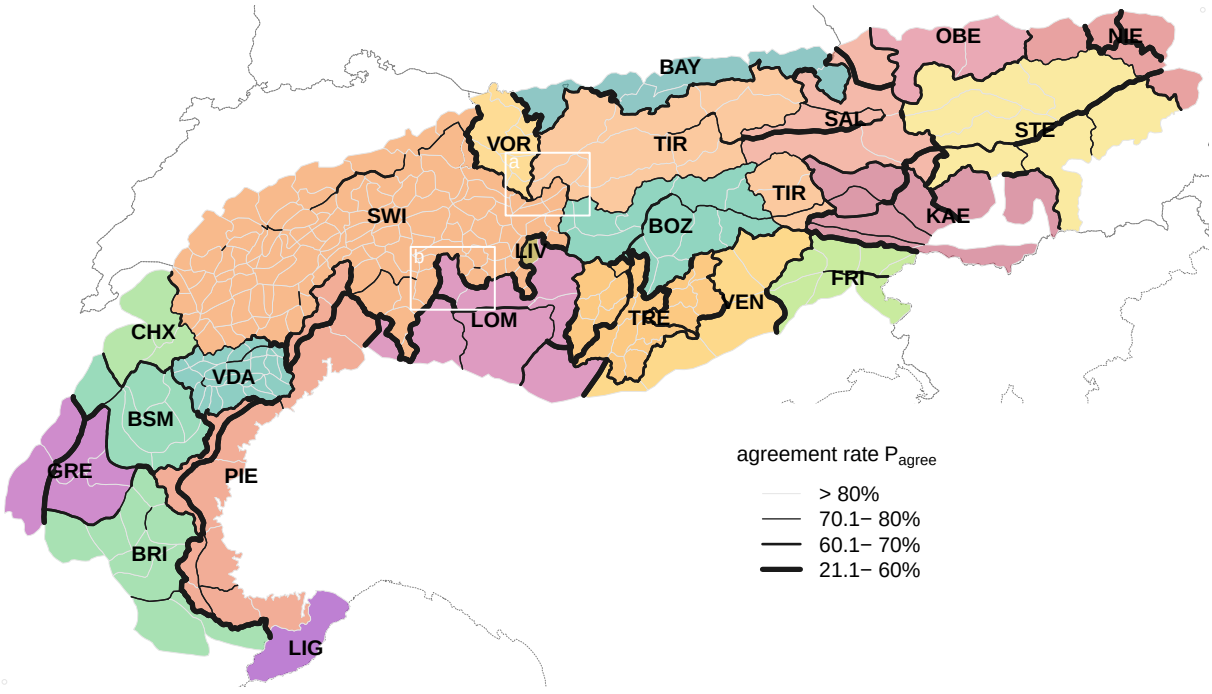


**Figure 5.** Map showing the individual forecast center domains in the European Alps (different colors, three-letter abbreviations see Table 2). The borders between warning regions are highlighted depending inversely on the agreement rate $P_{agree}$, with thicker lines corresponding to more frequent disagreements. The two white boxes (a, b) mark the two regions discussed in more detail in section 5.7.

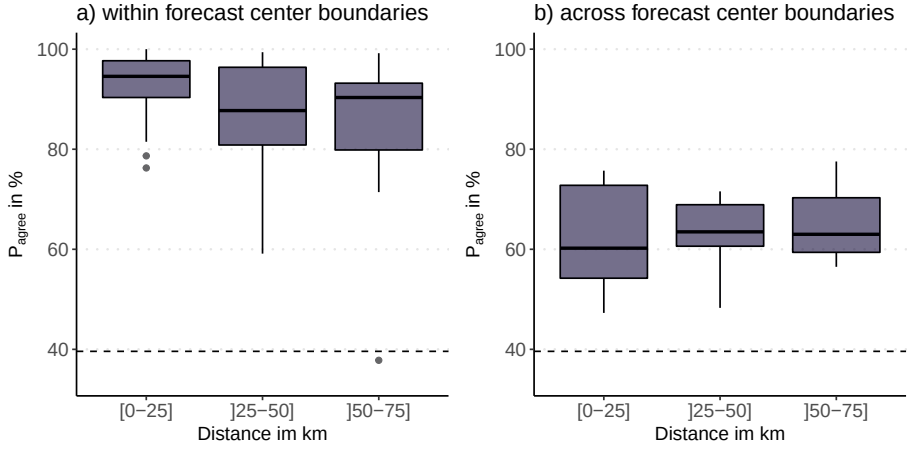


**Figure 6.** Boxplot showing the agreement rate ($P_{agree}$) for neighboring warning region pairs (a) within and (b) across forecast center boundaries, stratified by the distance between the center points of warning regions, with similar maximum elevation ($\Delta$ elevation $<$ 250 $\mathrm{m}$) and size (the size of the larger warning region is less than 1.5 times the size of the smaller warning region; N(within) = 108, N(across) = 37). The dashed line represents $P_{agree}$ when randomly drawing 10'000 danger levels for neighboring warning regions using the distributions shown in Fig. 4 (discussed in Section 6.2). The Boxplots show the median (bold line), the interquartile range (boxes), 1.5 times the interquartile range (whiskers) and outliers outside this range (dots).

lower than those for Friuli Venezia Giulia (FRI), Bourg-St-Maurice (BSM), Piemonte (PIE), Grenoble (GRE) and Briançon (BRI) (7.6%-12%). This can be partly attributed to more frequent occurrence of multi-day continuous periods with $D \geq 4$. Extended periods with $D \geq 4$ were comparably frequent in Briançon (BRI) or Piemonte (PIE) (more than 17% of these periods had a length of $\geq 3$ days), compared to Switzerland (SWI) and Chamonix (CHX) ($\geq$ 3 days: 4%). $P_{v.crit}$ in Briançon (BRI) was in many cases two or three classes higher compared to its immediate neighbors

in Italy (Piemonte (PIE), Liguria (LIG)), but also those in France (Bourg-St-Maurice (BSM), Grenoble (GRE)). The twelve regions with the highest $P_{v.crit}$ were clustered in the southwest of the Alps (9 in Briançon (BRI), 2 in Piemonte (PIE) and 1 in Grenoble (GRE), $P_{v.crit} \geq 9.8\%$, max = 15.3%).

$P_{v.crit}$ correlated very weakly with maximum elevation of a warning region ($\rho = 0.19$, $p < 0.01$). This correlation, however, was much stronger when exploring the proportion of days with $D \geq 3$ ($\rho = 0.7$, but also for $D = 3$ by itself ($\rho = 0.72$), see also Supplement S1).

### 5.4 Generally favorable avalanche situation D = 1

The proportion of days with a generally favorable avalanche situation $P_{favor}$ is

$$P_{favor} = \frac{N(D=1)}{N}. \quad (4)$$

Median $P_{favor}$ across the Alps was 5.3% (IQR: 3.4 - 13.8%), with two regions in Niederösterreich (NIE) having more than 50% of the forecasts with $D = 1$. The northern, southern and eastern rim of the Alps, generally regions with lower elevation (Fig. 3b), often have a larger proportion of days with favorable conditions (Fig. 7b). For regions with higher elevations, this proportion is lowest. This is also confirmed when correlating the maximum elevation of each warning region with $P_{favor}$ ($\rho$ = -0.75). In contrast, the correlation between $P_{favor}$ and the size of the warning regions is much weaker ($\rho$ = -0.26, $p < 0.001$).

Another obvious difference was the strong gradient between the eastern-most regions, where more than one third of the forecast period had generally favorable conditions, and those in the western and central parts of the Alps with comparably low values of $P_{favor}$.

### 5.5 Elevational gradients and temporal changes within forecast period

Different approaches are used to communicate elevational gradients in danger ratings (section 2.5). Forecast centers issuing two ratings - mostly in France, Austria and Bayern (BAY) - seldom indicated the highest hazard at lower elevations. This is in line with the correlations observed between the maximum elevation of a warning region and $P_{v.crit}$ (or $P_{favor}$, sections 5.3 and 5.4). The same danger rating was issued for all elevations by French forecast centers in two thirds of the forecasts, compared to 60% of the forecasts with an elevational gradient in Tirol (TIR) (Tab. 4).

All forecast centers, which were technically able to graphically communicate changes in danger level during the forecast period used this option. Most frequently, forecasts indicated no change during the forecast period (median 83%). Increasing danger levels ($D_{t2} > D_{t1}$) were communicated regularly by all the forecast centers (median 16%). However, the frequency varied considerably, between 26% in Vorarlberg (VOR) and less than 10% in Niederösterreich (NIE) and Oberösterreich (OBE, Tab. 5). Of particular note is Switzerland (SWI), the only warning service where increases in danger rating related to dry-snow avalanches were communicated exclusively in the textual danger description. A decrease in danger level during the forecast period was very rarely indicated (median 0.3%). Some forecast centers like Switzerland (SWI) never used this option. Notable exceptions were the forecasts by Vorarlberg (VOR) and Lombardia (LOM), where more than 6% of the forecasts indicated a decreasing danger rating within the forecast period.

**Table 4.** Elevational differences in danger rating with $D_{e1}$, the danger level above an indicated level, and $D_{e2}$, the danger rating below this elevation level. Example distributions are provided for some forecast centers.

| forecast center | $D_{e1} > D_{e2}$ | $D_{e1} = D_{e2}$ | $D_{e1} < D_{e2}$ |
|---|---|---|---|
| BRI, BSM, CHX, GRE | 32% | 67% | 0.9% |
| BAY | 45% | 48% | 7.2% |
| TIR | 60% | 35% | 4.6% |

**Table 5.** Temporal differences in danger rating within forecast period with $D_{t1}$, the danger rating valid for the first time step, and $D_{t2}$, for the second time-step. Example distributions are provided for some forecast centers.

| forecast center | $D_{t1} > D_{t2}$ | $D_{t1} = D_{t2}$ | $D_{t1} < D_{t2}$ |
|---|---|---|---|
| NIE, OBE | 0% | 95% | 5% |
| VOR | 13% | 61% | 26% |
| LOM | 6% | 72% | 22% |
| FRI, PIE | 0.2% | 74% | 25% |
| SWI | 0% | 87% | 13%* |
| BRI, BSM, CHX, GRE | 0.9% | 84% | 15% |

*For Switzerland, the proportion of changing danger ratings which are exclusively communicated in the danger description is 2.7%.

### 5.6 Size of the warning regions, $P_{v.crit}$ and spatial variation in danger level

As outlined in section 2.4 and shown in Figure 3, varying spatial scales and approaches are used to produce the forecast, and communicate a danger level. One of these approaches relies on a comparably fine spatial resolution of the warning regions in the bulletin production process, as is the case in Valle d'Aosta (VDA, Italy), Switzerland (SWI) and Trentino (TRE).

The forecast center VDA uses 26 warning regions (median size 130 $\mathrm{km}^2$, Tab. 2, Fig. 3). Each of these regions belongs to one of four larger snow-climate regions (me-

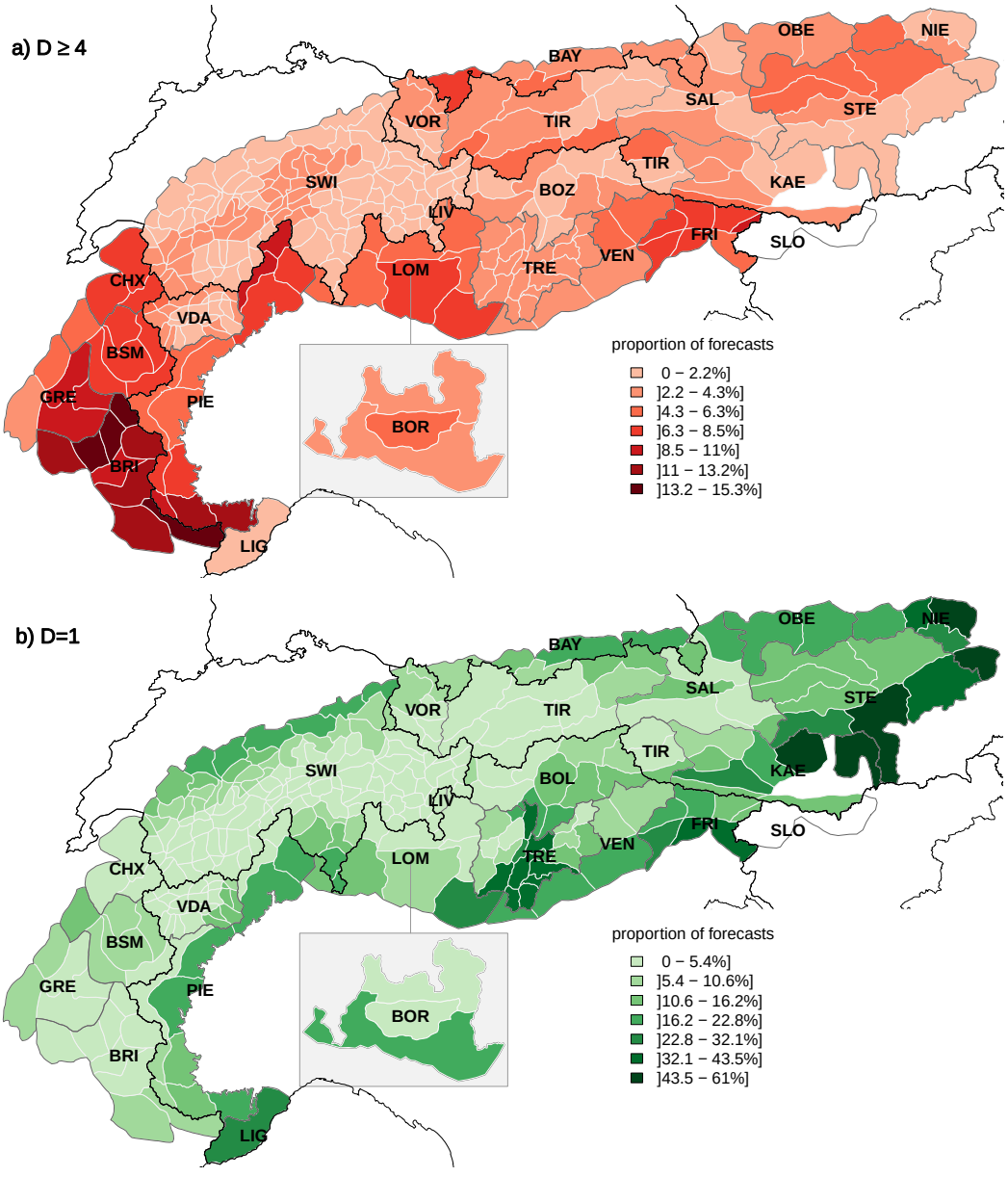


**Figure 7.** Map showing the European Alps with (a) the proportion of days with forecast very critical conditions ($P_{v.crit}$, $D_{max} \geq 4$) and (b) with a forecast danger level 1 ($P_{favor}$, $D_{max} = 1$). The color shading of the individual warning regions (white borders) corresponds to the values of $P_{v.crit}$ and $P_{favor}$, respectively. Forecast centers are labeled according to Table 2 and marked with dark grey polygon borders, national borders with black lines. To visualize the (at least partially) overlapping forecast regions in the Italian region of Lombardia, LIV is superposed onto parts of LOM, while BOR is placed as inset to the south of LOM. Thresholds for the color classes were defined using the Fisher-Jenks algorithm minimizing within-class variation (Slocum et al. (2005); R-package *classInt* Bivand (2017)).

dian size 815 $\mathrm{km}^2$, Burelli et al. (2016, p. 27)). In Switzerland, the forecaster aggregates the 117 warning regions in the Swiss Alps (median size 180 $\mathrm{km}^2$) to (generally) three to five regions with the same danger description (with an average size per aggregated region of 5000 - 7000 $\mathrm{km}^2$; Ruesch et al., 2013; Techel and Schweizer, 2017). Similar to VDA, each of the Swiss warning regions can be linked to a higher-order spatial hierarchy (SLF, 2015, p. 41) [3]. In either case, these predefined regional aggregations are not of great importance anymore in the communication of a regional danger level, due to the flexibility in which the forecaster can assign danger ratings to regions (VDA)

[3] As an example, the warning region «1121 - Freiburger Alpen» belongs at its highest hierarchy level to the snow-climate region «1 - western part of the Northern flank of the Alps».

or aggregate regions (SWI). However, here we use these spatial hierarchy-levels - three for VDA and four for SWI[4] - to explore the variability of the forecast danger level within regions of increasing size and the potential implication on summary statistics like the proportion of the most critical forecasts ($P_{v.crit}$, Section 5.3).

As shown in Tab. 6, the larger a region, the higher the variability within these regions (more than one danger level forecast). In other words, a forecaster would not have been able to communicate the spatial variability in danger levels without describing these in text form if warning regions were five times larger (about 800 $\mathrm{km}^2$, corresponding to the median size in Niederösterreich (NIE) or in France) in about 15% of the forecasts, as compared to the currently implemented spatial resolution. Assuming even larger warning regions at the communication level, 3300 $\mathrm{km}^2$, for instance when considering VDA as one single region, or the seven snow-climate regions in SWI, and communicating a single danger rating only, would have resulted in about half of the forecasts not reflecting the spatial variability within the respective region.

This shows that variations in the expected avalanche hazard at spatial scales lower than the size of the spatial units used in the production and communication of the forecast are to be expected, particularly if regions are large. In these situations, a forecaster must decide whether to communicate the highest expected danger level, regardless of its spatial extent, or the danger level representative for the largest part of a region. Note that currently the EADS lacks a definition in that respect. Taking the proportion of forecasts with very critical conditions $P_{v.crit}$ shows that communicating the highest danger level within a region $P_{v.crit}$(max) increases the absolute values of $P_{v.crit}$ (Tab. 6). Communicating the spatially most widespread danger rating instead ($P_{v.crit}$(mean)), has relatively little influence for smaller regions, but reduces $P_{v.crit}$ values significantly for the largest-size regions (Tab. 6).

At the current spatial resolution, $P_{v.crit}$ values for SWI and VDA are comparable, particularly along their joint border (Fig. 7a). However, $P_{v.crit}$(max) values at the first-order aggregation are already considerably higher for VDA, and rather similar to those in neighboring warning regions in Chamonix (CHX), Bourg-St-Maurice (BSM) or Piemonte (PIE).

### 5.7 Case studies

To make the results more tangible, we present two case studies (Fig. 8):

The *Silvretta* mountain range, at the border between Austria (Vorarlberg (VOR) and Tirol (TIR)) in the North and Switzerland (SWI) in the South (Fig. 8a) is split into six warning regions, all including *Silvretta*, and/or *Samnaun* in their region name. These have similar maximum elevations (between 3200 and 3340 m), but differ in size (SWI $\leq$ 180 $\mathrm{km}^2$, TIR 490 $\mathrm{km}^2$). According to Schwarb et al. (2001), there is a precipitation gradient during the three winter months December to February with total precipitation amounts decreasing from about 250 - 300 $\mathrm{mm}$ (in VOR) to about 150 - 200 $\mathrm{mm}$ in the Eastern most regions in TIR.

The agreement rate is high between the Swiss *Silvretta* regions (93%), but considerably lower across forecast center boundaries (SWI - TIR 73%, SWI - VOR 64%). Note further, that between the Swiss *Silvretta* and *Samnaun* $P_{agree}$ equals 100%. Additionally, there is a significant bias present between SWI and its two Austrian neighbors ($p < 0.001$), with the danger level in Switzerland being lower more often than higher. In contrast, despite a low agreement rate (67%) there is no significant bias between TIR and VOR, implying that differences in forecast avalanche danger are balanced. Note further, that $P_{agree}$ between VOR and its neighbors in SWI or TIR is 5 to 10% higher when considering $D_{morning}$ rather than $D_{max}$. Danger level 4 was least often forecast in the Swiss warning regions ($P_{v.crit} < 1.2\%$) and most often in the largest of the five regions, in Tirol (4.7%). In comparison, $D = 1$ was forecast between 2.4% in Tirol and 4.7% in the two western-most regions in Vorarlberg and Switzerland.

Turning to a location south of the main Alpine divide, where the Italian *Retiche occidentale* warning region in Lombardia (LOM, size 510 $\mathrm{km}^2$, elevation 3200 $\mathrm{m}$) lies embedded between three Swiss warning regions (SWI, size 120 - 370 $\mathrm{km}^2$, elevation 2900 - 3300 $\mathrm{m}$). It is an area, which receives most precipitation from southerly air currents. Winter precipitation is generally more abundant in the Southwest (200 - 250 $\mathrm{mm}$) compared to the North and East of these regions (150 - 200 $\mathrm{mm}$, Schwarb et al., 2001). This pattern is more pronounced in spring (March - May). The agreement rate between the three Swiss warning regions was between 79% and 90%, despite them being sometimes separated by the Lombardian warning region. The agreement rate between the Swiss and Lombardian region ranged between 47% and 59%. The bias was very pronounced with Swiss forecasts often being lower than the ones in LOM ($p < 0.001$). This also shows when comparing $P_{v.crit}$ ($P_{v.crit}$(LOM) $=$ 5.2% vs. $P_{v.crit}$(SWI) $<$ 1.2%) or $P_{favor}$ ($P_{favor}$(LOM) $=$ 1.8% vs. $P_{favor}$(SWI) $>$ 3.8%).

## 6 Discussion

We explored spatial consistency and bias using published forecast avalanche danger levels by using a comparably large number of real forecasts rather than a small number of hypothetical scenarios, as in the experiment conducted

[4] no higher hierarchy exists for the warning regions in TRE

**Table 6.** Variability in danger ratings and the proportion of forecasts with danger levels *4-High* or *5-Very High* ($P_{v.crit}$) assuming different aggregation levels as the given spatial resolution for danger level communication. The aggregation level *none* indicates the currently used spatial resolution. The aggregated median size and number (N) of regions within the forecast domain are indicated. $P_{v.crit}$(max) assumes the communication of the highest danger rating per region, and $P_{v.crit}$(mean) the spatially most relevant danger rating.

| forecast center | aggregation | size (km$^2$) | N | 1 rating | 2 ratings | $\geq$ 3 ratings | $P_{v.crit}$(max) | $P_{v.crit}$(mean) |
|---|---|---|---|---|---|---|---|---|
| VDA | none | 130 | 26 | 100% | - | - | 2.3% | 2.3% |
| | first-order | 815 | 4 | 83% | 17% | 0.3% | 3.7% | 2.3% |
| | second-order* | 3300 | 1 | 56% | 39% | 5% | 6.8% | 0.7% |
| SWI | none | 180 | 117 | 100% | - | - | 1.3% | 1.3% |
| | first-order | 740 | 35 | 85% | 15% | 0.3% | 1.6% | 1.3% |
| | second-order | 1740 | 17 | 71% | 28% | 1.1% | 2.3% | 1.3% |
| | third-order | 3260 | 7 | 53% | 44% | 2.9% | 3.1% | 1% |

* - considering the entire VDA forecast domain as one region

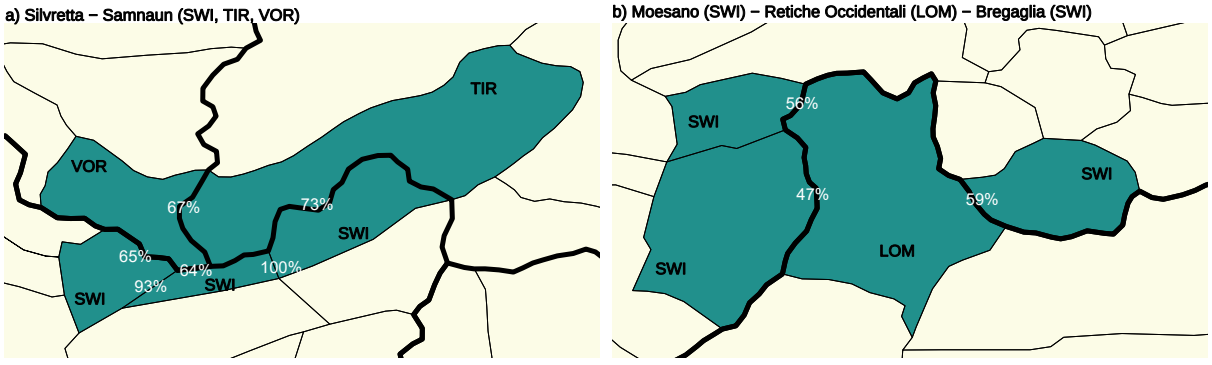


**Figure 8.** Example regions: a) Silvretta mountain range with the *Silvretta* warning regions in Vorarlberg (VOR) and Tirol (TIR) and three Swiss warning regions (SWI, from west to east: *Western Silvretta*, *Eastern Silvretta* and *Samnaun*). b) the *Rhetiche Occidentali* warning region (forecast center Lombardia (LOM)) and the three Swiss warning regions *Alto Moesano*, *Basso Moeasano* and *Bregaglia*. Here, the main Alpine divide runs right to the north of the dark-colored regions.
The percentage values show the agreement rate between warning regions ($D_{max}$). The maps show an area of 83 by 45 km. The location of these two example regions in the Alps is marked in Fig. 5.

by Lazar et al. (2016). However, using actual forecasts in such a diverse setting as the European Alps, comes at the cost of many confounding factors. Differences between forecast centers in the forecast production and danger level communication required us to make some assumptions prior to data analysis.

In this discussion, we first summarize the main quantitative findings, which we then put into perspective given the data (Section 6.1) and our methodology (Section 6.2). Furthermore, we discuss sources for inconsistencies and bias (Section 6.3) and potential implications to forecast users (Sect. 6.4).

The main results are:

- The agreement rate $P_{agree}$ was significantly lower across national and forecast center boundaries (about 60%), compared to within forecast center boundaries (about 90%, Figures 5 and 6).
- Significant bias was often observed across national and forecast center boundaries, with several forecast centers showing systematic differences towards lower (or higher) danger levels than their neighbors (Section 5.2).
- The proportion of forecasts with danger levels *4-High* and *5-Very High* showed considerable spatial variability (Fig. 7a), with pronounced differences across some forecast center boundaries, and was influenced by the size of warning regions (Section 5.6).

### 6.1 Dataset: four winter seasons

We explored avalanche forecasts published during four winter seasons (477 forecast days). These included the 2011/2012 winter with extended periods of heavy snowfalls affecting particularly the regions north of the main Alpine Divide (Northern French Alps, large parts of Switzerland and Austria, Bavarian Alps; Coléou, 2012;

ÖLWD, 2012; Techel et al., 2013), but also the 2013/2014 winter, which was one of the snowiest winters on record in the Southern Alps (Italy, southern parts of Switzerland; Goetz, 2014; ÖLWD, 2014; Techel et al., 2015a; Valt and Cianfarra, 2014). These two winters, or removing one of them during data analysis, had an effect particularly on the absolute values of the proportion of forecasts with $D \geq 4$ ($P_{v.crit}$), while the overall rank order remained comparably similar, regardless of which subset was analyzed Supplement S2 Removing individual winters also had no significant influence on the agreement rate ($P_{agree}$) or bias ($B_{ij}$) between neighboring warning regions. By comparing with long-term statistics of forecast danger levels (e.g. France, Switzerland, Steiermark; Mansiot, 2016; Techel et al., 2013; Zenkl, 2016), we conclude that our data are generally representative and the four years analyzed cover a typical range of conditions encountered in the European Alps.

## 6.2 Methodology

Danger levels were communicated in different ways in the forecasts (Section 2.5). Therefore, we generalized by defining two data subsets which could be applied to most forecast products: $D_{max}$, describing the highest danger rating within a forecast period, valid for (part of) the day and the most exposed elevations, and $D_{morning}$, where we assumed that time step 1 generally referred to the morning, and time step 2 to the afternoon.

Using $D_{max}$ or $D_{morning}$ for analysis influenced absolute values of $P_{v.crit}$ (Sect. 5.3), but less the rank order, and had little influence on $P_{agree}$ or $B_{ij}$ (Sect. 5.2).

We introduced $P_{agree}$ as a measure of spatial consistency (or correlation). As shown in Fig. 4, on four of five days $D = 2$ or $D = 3$ were forecast. Thus, by chance alone, a minimal agreement rate can be expected. We estimated this minimal agreement rate by simulating 10'000 danger levels for two neighboring regions using the danger level distributions shown in Fig. 4. Doing so we obtained values of $P_{agree} = 40\%$ for $D_{max}$ and $P_{agree} = 36\%$ for $D_{morning}$. Thus, levels of agreement reported in this paper, and in any future work, should be compared with a minimal agreement rate based on realistic values derived from observed danger level distributions.

Similarly, total agreement ($P_{agree}$ = 100%) between neighboring regions implies that subdivisions may be superfluous. Nonetheless, we found 100% agreement for a total of 14 warning region pairs in Switzerland, Italy and Austria. To confirm whether this agreement indicates regions which could be merged would require further investigation as to, for example, the nature of typical avalanche problems found, and not only the forecast danger levels.

The spatial resolution of the warning regions (Tab. 2, Fig. 3a), and how these are used in the communication of the forecasts, varied greatly between forecast centers. As we have shown for the forecasts in Switzerland (SWI) and Valle d'Aosta (VDA; Section 5.6), this may in turn influence the danger rating communicated to the public. As a consequence, it has an impact on all summary statistics, most notably $P_{v.crit}$ and $B_{ij}$.

We explored a mix of forecasts for the day of publication, the following day, or even the day after. However, forecast accuracy generally decreases with lead time (Jamieson et al., 2008; Statham et al., 2018b). Forecast accuracy may also vary within forecast center domains, as shown by Techel and Schweizer (2017) for the case of Switzerland. We suspect that these may affect primarily the agreement rate $P_{agree}$, except if the forecast bias differs temporally or spatially.

Within forecast center domains, differences in the frequency of the danger levels, the agreement rate $P_{agree}$, or the bias $B_{ij}$ may indicate differences in snow avalanche climate. In all other situations, that is to say when looking at differences between forecast centers, operational constraints must be considered as much as snow-climate, when exploring consistency and bias.

## 6.3 Understanding differences between avalanche warning regions

Our aims in exploring spatial consistency and bias were threefold: firstly to investigate whether differences existed between forecasting centers, secondly to understand potential factors influencing these biases, and finally to consider the influence of these biases on forecast users. Our results clearly demonstrate that spatial inconsistencies and biases exist, above all across forecast center boundaries. In the following we briefly discuss three possible reasons for such differences, two of which suggest limitations in current forecasting approaches.

The size of the warning regions differed considerably between forecast centers (Fig. 3, Tab. 2) and had an impact on the issued danger level in general, particularly on $P_{v.crit}$ (Section 5.6). Coarser spatial resolutions of warning regions lead not only to more forecasts with higher danger levels, but also increase variability within warning regions. Such variability cannot be captured with a single value and thus, though it may be expressed within the forecast text, is ignored by our approach. Since differences in warning region size were correlated with both bias and agreement rate, we recommend exploring whether more heterogeneous warning regions - from an avalanche winter regime perspective - might be divided into smaller ones to reduce such bias. We also found correlations between avalanche danger levels, bias, agreement rate and elevation. While higher elevations and higher avalanche dangers are often associated with one another, we suggest the relationship between bias and elevation may result from different ways of communicating avalanche danger for a warning

region. In particular, the EADS does not specify whether the highest, or the spatially most representative danger level should be communicated for a warning region. We therefore suggest that the EAWS consider whether being more specific in defining how avalanche danger should be assigned to a warning region may reduce bias.

This lack of specificity in the EADS with respect to avalanche danger is an example of potential differences in the application of the EADS in different forecast centers, which may in turn explain some aspects of inconsistency and bias. Simply put, forecasters must assign a categorical value to a complex forecast, which typically also contains uncertainty. This assignment of an avalanche danger level is not only influenced by conditions, but may also emerge from cultural differences in forecasting practices (McClung, 2000; Greene et al., 2006; Lazar et al., 2016) and explicit or implicit internalization by forecasters of the use and implication of danger levels by local, regional and national risk management authorities. The need to increase consistency in the application of the EADS has been recognized. Efforts made by the EAWS include improvements in the *EAWS matrix*, a tool assisting forecasters in assigning danger levels (Müller et al., 2016; EAWS, 2017a) and the provision of clear definitions of key contributing factors, such as the distribution of dangerous locations and the likelihood of avalanche release. Nonetheless, it is important to recognize that even if the EAWS strive to harmonize practices and production, externalities such as the consequences of danger levels for users, and the perception of forecasters of this impact, may alter the homogeneity of the product. Furthermore, as observed by LaChapelle (1980) and summarized very recently by Statham et al. (2018a), avalanche forecasts are produced by a forecaster making subjective judgments based on the available data and evidence. Reducing these forecasts to a categorical value neither removes the subjectivity in the process nor does it allow the forecaster to communicate uncertainty.

A third possible reason for differences between warning regions lies not in bias or inconsistency in the use of the EADS, but rather in real differences in the avalanche winter regime (Haegeli and McClung, 2007). Many of the warning region boundaries, especially along national borders, follow the main Alpine divide, which also serves as a main weather divide. Where large differences in avalanche winter regime are observed, a lower correlation in danger ratings would therefore be expected. However, we relied exclusively on forecast danger levels and cannot compare the agreement rate or bias with differences in avalanche winter regime. This is an important limitation in our study. Incorporating avalanche winter regimes in this study, and/or typical avalanche problems - if these were used consistently, would clearly be beneficial for the interpretation of our findings. Such an analysis would require, besides meteorological data, a common database containing snow structure and avalanche information for the entire Alpine mountain range, as already exists for the US and Canada (Mock and Birkeland, 2000; Haegeli and McClung, 2007; Shandro and Haegeli, 2018).

### 6.4 Inconsistencies: implications for forecast users

A final key question is the implications of the potential spatial inconsistencies and biases in the use of danger levels for forecast users. Even though there may be good reasons for such differences, such as the difference in size of warning regions and therefore a need to communicate different information, users are unlikely to appreciate or understand such nuances.

Regional avalanche forecasts are considered an important source of information for backcountry users, particularly during the planning stage, but also on the day of the tour (Winkler and Techel, 2014; LWD Steiermark, 2015; Baker and McGee, 2016). A key advantage of the introduction of the EADS in 1993 was seen as the provision of consistent information across the European Alps (Meister, 1995). Forecast danger level has been shown to be the part of the forecast most known and used in the Alps (Winkler and Techel, 2014; LWD Steiermark, 2015; Procter et al., 2014), influencing backcountry destinations (Techel et al., 2015b) and local decision-making by recreationists (Furman et al., 2010). Many users of avalanche forecasts are typically active within warning regions where forecasts are produced by a single regional avalanche forecast center (e.g. in Voralberg (VOR) or Tirol (TIR)). Such users are likely to become accustomed and calibrated to «their» forecast. Thus, issues are likely to arise when users travel from one forecast center domain to another. For instance, a frequent user of French forecasts traveling to Switzerland may experience some Swiss forecasts with D = 3 as a missed alarm, while the opposite may happen when a Swiss user recreates in France. In both cases this reduces the credibility of the forecasts, as they are perceived to be less accurate (Williams, 1980). We suggest that harmonization efforts should therefore focus not only on the product - an avalanche forecast - but how this product is used and interpreted by different users and their requirements (Murphy, 1993).

## 7 Conclusions

In this study, we explored the avalanche forecast products, and specifically the forecast danger level during four years with 477 forecast days from 23 forecast centers in the European Alps. For the first time,

- (i) we qualitatively described the operational constraints in the production and communication of danger level in avalanche forecast products in the Alps,

- (ii) we developed a methodology to explore spatial consistency and bias in avalanche forecasts,
- (iii) we quantified spatial consistency and bias in forecast danger levels, given operational constraints and the selected methods, and
- (iv) we discuss the implications of spatial consistency and bias for forecasting and forecast users.

We noted considerable differences in the operational constraints associated with forecast products. Most notably the spatial resolution of the warning regions underlying the forecasts had an impact on biases observed and the agreement rate, but also limits at what spatial scale a regional danger level can be communicated in map products. Furthermore, we detected discrepancies in the use of the higher danger levels, as well as a comparably large proportion of forecasts with different danger levels across forecast center boundaries. These findings indicate a need to further harmonize the production process and communication of avalanche forecast products, not just across the Alps but throughout Europe. Harmonization should consider

- (i) similar approaches regarding the size of warning regions and their aggregation, with a preference towards using a finer spatial resolution,
- (ii) focusing not only on forecast products, but also user requirements, and
- (iii) the consistent use of EADS by incorporating the *EAWS-Matrix*, and further developments, and developing a consistent workflow, similar to the approach suggested by Statham et al. (2018a) into the production process.

To carry out our study we had to collect and harmonize data across the Alps. We recommend a development of a centralized system for collecting data, which would enable further studies of forecast properties in the future.

*Data availability.* The data will be made freely available on the data portal www.enividat.ch

*Author contributions.* Frank Techel coordinated and designed this collaborative study. All co-authors provided repeatedly in-depth feedback regarding previous versions of this manuscript. Authors are listed alphabetically.

*Competing interests.* No competing interests are present.

*Acknowledgements.* This study would not have been possible without the contributions of numerous avalanche forecasters and researchers, who provided data and/or feedback on the manuscript (in alphabetical order): Igor Chiambretti, Jean-Louis Dumas, Mattia Faletto, Thomas Feistl, Elisabeth Hafner, Fabiano Monti, Patrick Nairz, Bernhard Niedermoser, Andreas Pecl, Evelyne Pougatch, Jürg Schweizer, Luca Silvestri, Florian Stifter, Thomas Stucki, Arnold Studeregger, Lukas Rastner, Mauro Valt, Romeo Vicenzo, Alec van Herwijnen, Gernot Zenkl.
We greatly appreciate the detailed and constructive comments by the three reviewers Karsten Müller, Rune Engeset and Karl Birkeland.
The statistical analysis was performed using the open-source software *R, version 3.4.3* (R Core Team, 2017).

## References

Baker, J. and McGee, T.: Backcountry snowmobilers' avalanche-related information-seeking and preparedness behaviors, Society & Natural Resources, 29, 345 – 356, doi:10.1080/08941920.2015.1103387, 2016.

Ballou, D. and Pazer, H.: Modeling completeness versus consistency tradeoffs in information decision contexts, IEEE Transactions on Knowledge and Data Engineering, 15, 240–243, doi:10.1109/TKDE.2003.1161595, 2003.

Bivand, R.: classInt: Choose univariate class intervals, https://CRAN.R-project.org/package=classInt, r package version 0.1-24, 2017.

Bivand, R. and Piras, G.: Comparing implementations of estimation methods for spatial econometrics, Journal of Statistical Software, 63, 1–36, http://www.jstatsoft.org/v63/i18/, 2015.

Bivand, R., Pebesma, E., and Gómez-Rubio, V.: Applied spatial data analysis with R, Springer Science + Business Media New York 2013, 2 edn., doi:10.1007/978-1-4614-7618-4, 2013.

Bovee, M., Srivastava, R., and Mak, B.: A conceptual framework and belief-function approach to assessing overall information quality, International Journal of Intelligent Systems, 18, 51–74, doi:10.1002/int.10074, 2003.

Burelli, G., Ceaglio, E., Contri, G., Debernardi, A., Frigo, B., and Pivot, S.: Rendiconto nivometeorologico. Inverno 2011-2012, Tech. rep., Regione Autonome Valle d'Aosta e Fondazione Montagna sicura, 184 p., 2012.

Burelli, G., Ceaglio, E., Contri, G., Debernardi, A., and Pivot, S.: Rendiconto nivometeorologico. Inverno 2015-2016, Tech. rep., Regione Autonome Valle d'Aosta e Fondazione Montagna sicura, 152 p., 2016.

Coléou, C.: Enneigement dans les massifs françaises dans l'hiver 2011-2012, Neige et Avalanche, 139, 28 – 29, 2012.

Dale, M. and Fortin, M.-J.: Spatial analysis: a guide for ecologists, Cambridge University Press, 2 edn., 2014.

EAWS: Bavarian matrix, http://www.avalanches.org/eaws/en/main_layer.php?layer=basics&id=5, 2005.

EAWS: EAWS Matrix, Tech. rep., http://www.avalanches.org/eaws/en/main_layer.php?layer=basics&id=4, last access: 2017/11/12, 2017a.

EAWS: Content and structure of public avalanche bulletins, Tech. rep., https://lawine.tirol.gv.at/data/produkte/basics/ContentAndStructureAvalancheBulletin.pdf, last access: 2017/06/01, 2017b.

EAWS: Memorandum of understanding for the European Avalanche Warning Services (EAWS), Tech. rep., https://lawine.tirol.gv.at/data/eaws/MoU_EAWS.pdf, last access: 02/09/2018, 2017c.

EAWS: Typical avalanche problems, Tech. rep., https://lawine.tirol.gv.at/data/eaws/typical_problems/EAWS_avalanche_problems_EN.pdf, approved by General Assembly of EAWS, Munich, 2017; last access: 01/12/2017, 2017d.

EAWS: European Avalanche Danger Scale (2018/19), http://www.avalanches.org/eaws/en/main_layer.php?layer=basics&id=1, last access 06/10/2016, 2018.

Elder, K. and Armstrong, B.: A quantitative approach for verifying avalanche hazard ratings, in: Symposium at Davos 1986 on Avalanche Formation, Movement and Effects, vol. 162 of *International Association of Hydrological Sciences Publication*, pp. 593 – 603, 1987.

ESRI: ArcGIS online basemap: world topo map, 2017.

Föhn, P. and Schweizer, J.: Verification of avalanche danger with respect to avalanche forecasting, in: Les apports de la recherche scientifique à la sécurité neige, glace et avalanche. Actes de Colloque, Chamonix, vol. 162, pp. 151–156, Association Nationale pour l'Étude de la Neige et des Avalanches (ANENA), 1995.

Furman, N., Shooter, W., and Schumann, S.: The roles of heuristics, avalanche forecast, and risk propensity in the decision making of backcountry skiers, Leisure Sciences, 32, 453–469, doi:10.1080/01490400.2010.510967, 2010.

Giraud, G., Lafeuille, J., and Pahaut, E.: Evaluation de la qualité de la prévision du risque d'avalanche, Int. Ass. Hydrol. Sci. Publ., 162, 583–591, 1987.

Goetz, D.: Bilan nivo-météorologique de l'hiver 2013-2014, Neige et Avalanche, 147, 12 – 14, 2014.

Greene, E., Wiesinger, T., Birkeland, K., Coléou, C., Jones, A., and Statham, G.: Fatal avalanche accidents and forecasted danger level: patterns in the United States, Canada, Switzerland and France, in: Proceedings ISSW 2006. International Snow Science Workshop, Telluride, Co., pp. 640–649, 2006.

Haegeli, P. and McClung, D.: Expanding the snow-climate classification with avalanche-relevant information: initial description of avalanche winter regimes for southwestern Canada, J. Glaciol., 53, 266–276, 2007.

Hijmans, R.: raster: Geographic data analysis and modeling, https://CRAN.R-project.org/package=raster, r package version 2.5-8, 2016.

Jamieson, B., Campbell, C., and Jones, A.: Verification of Canadian avalanche bulletins including spatial and temporal scale effects, Cold Reg. Sci. Technol., 51, 204–213, doi:10.1016/j.coldregions.2007.03.012, 2008.

Jarvis, A., Reuter, H., Nelson, A., and Guevara, E.: Hole-filled seamless SRTM data V4, Tech. rep., International Centre for Tropical Agriculture (CIAT), http://srtm.csi.cgiar.org, 2008.

LaChapelle, E.: The fundamental process in conventional avalanche forecasting, Journal of Glaciology, 26, 75–84, 1980.

Lazar, B., Trautmann, S., Cooperstein, M., Greene, E., and Birkeland, K.: North American avalanche danger scale: Do backcountry forecasters apply it consistently?, in: Proceedings ISSW 2016. International Snow Science Workshop, Breckenridge, Co., pp. 457 – 465, 2016.

LWD Steiermark: Ergebnisse der Online-Umfrage des LWD Steiermark 2015, 2015.

Mansiot, O.: Indices de risque 2 et 3: utilise-t-on correctement les informations du bulletin d'estimation du risque d'avalanche?, Neige et Avalanches, 155, 8–10, 2016.

Marazzi, S.: Atlante orografico delle Alpi: SOIUSA - suddivisione orografica internazionale unificata del Sistema Alpino, Pavone Canavese: Priuli & Verlucca, 2005.

McClung, D.: Predictions in avalanche forecasting, Annals of Glaciology, 31, 377–381–, 2000.

Meister, R.: Country-wide avalanche warning in Switzerland, in: Proceedings ISSW 1994. International Snow Science Workshop 1994, Snowbird, UT, pp. 58–71, 1995.

Météo France: Guide avalanche, Météo France, Saint-Mandé Cedex, édition 2012-2013 edn., 2012.

Mock, C. and Birkeland, K.: Snow avalanche climatology of the Western United States mountain ranges, Bulletin of the American Meteorological Society, 81, 2367 – 2392, doi:10.1175/1520-0477(2000)081<2367:SACOTW>2.3.CO;2, 2000.

Müller, K., Stucki, T., Mitterer, C., Nairz, P., Konetschny, H., Feistl, T., Coléou, C., Berbenni, F., and Chiambretti, I.: Towards an improved European auxiliary matrix for asessing avalanche danger levels, in: Proceedings ISSW 2016. International Snow Science Workshop, Breckenridge, Co., USA, 2016.

Munter, W.: 3x3 Lawinen, Agentur Pohl und Schellhammer, Garmisch-Partenkirchen, 1st edn., 1997.

Murphy, A. H.: What is a good forecast? An essay on the nature of goodness in weather forecasting, Weather and Forecasting, 8, 281–293, doi:10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2, 1993.

ÖLWD: Saisonbericht der österrreichischen Lawinenwarndienste 2011/2012, Arbeitsgemeinschaft der österreichischen Lawinenwarndienste, 180 p., 2012.

ÖLWD: Saisonbericht der österrreichischen Lawinenwarndienste 2013/2014, Arbeitsgemeinschaft der österreichischen Lawinenwarndienste, 236 p., 2014.

ÖLWD: Saisonbericht der österrreichischen Lawinenwarndienste 2014/2015, Arbeitsgemeinschaft der österreichischen Lawinenwarndienste, 272 p., 2015.

Pahaut, E. and Bolognesi, R.: Prévisions régionale et locale du risque d'avalanches, in: Guide Neige et avalanche. Connaissances, Pratiques, & sécurité, edited by Ancey, C., ISSN 2-85744-797-3, chap. 7, pp. 161–178, Édisud, Aix-en-Provence, France, 3 edn., electronic version, based on second edition, 2003.

Procter, E., Strapazzon, G., Dal Cappello, T., Castlunger, L., Staffler, H., and Brugger, H.: Adherence of backcountry winter recreationists to avalanche prevention and safety practices in northern Italy, Scand. J. Med. Sci. Sports, 24, doi:10.1111/sms.12094, 2014.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, last updated: June 2017, 2017.

Ruesch, M., Egloff, A., Gerber, M., Weiss, G., and Winkler, K.: The software behind the interactive display of the Swiss avalanche bulletin, in: Proceedings ISSW 2013. International

Snow Science Workshop, Grenoble – Chamonix-Mont Blanc, France, pp. 406–412, 2013.

Schwarb, M., Frei, C., Schär, C., and Daly, C.: Hydrologischer Atlas der Schweiz, chap. 2.7 - Mean Seasonal Precipitation throughout the European Alps 1971–1990, Geographisches Institut der Universität Bern – Hydrologie, 2001.

Schweizer, J., Kronholm, K., and Wiesinger, T.: Verification of regional snowpack stability and avalanche danger, Cold Reg. Sci. Technol., 37, 277–288, doi:10.1016/S0165-232X(03)00070-3, 2003.

Schweizer, J., Mitterer, C., Techel, F., Stoffel, A., and Reuter, B.: Quantifiying the obvious: the avalanche danger level, in: Proceedings ISSW 2018. International Snow Science Workshop, Innsbruck, Austria., 2018.

Shandro, B. and Haegeli, P.: Characterizing the nature and variability of avalanche hazard in western Canada, Natural Hazards and Earth System Sciences, 18, 1141–1158, doi:10.5194/nhess-18-1141-2018, 2018.

SLF: Interpretationshilfe zum Lawinenbulletin des Eidgenössischen Institutes für Schnee- und Lawinenforschung Weissfluhjoch, Davos, in: Mitteilung des Eidgenössischen Institutes für Schnee- und Lawinenforschung, vol. 49, p. 24, Eidgenössischen Institutes für Schnee- und Lawinenforschung, 1993.

SLF: Avalanche bulletins and other products. Interpretation guide. Edition 2015, WSL Institute for Snow and Avalanche Research SLF, Davos, http://www.slf.ch/lawineninfo/zusatzinfos/interpretationshilfe/interpretationshilfe_e.pdf, 16th revised edition, 50p., 2015.

Slocum, T., McMaster, R., Kessler, F., and Howard, H.: Thematic cartography and geographic visualization, Prentice Hall Series in Geographic Information Science, Pearson/Prentice Hall, Upper Saddle River, NJ, 2 edn., 2005.

SRTM: SRTM data V4, Website, http://www.cgiar-csi.org, download 19/04/2017, 2017.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J.: A conceptual model of avalanche hazard, Natural Hazards, 90, 663 – 691, doi:10.1007/s11069-017-3070-5, 2018a.

Statham, G., Holeczi, S., and Shandro, B.: Consistency and accuracy of public avalanche forecasts in Western Canada, in: Proceedings ISSW 2018. Inteernational Snow Science Workshop, Innsbruck, Austria., 2018b.

Techel, F. and Schweizer, J.: On using local avalanche danger level estimates for regional forecast verification, Cold Regions Science and Technology, 144, 52 – 62, doi:10.1016/j.coldregions.2017.07.012, 2017.

Techel, F., Pielmeier, C., Darms, G., Teich, M., and Margreth, S.: Schnee und Lawinen in den Schweizer Alpen. Hydrologisches Jahr 2011/12, WSL-Institut für Schnee- und Lawinenforschung SLF Davos: 118 pages (WSL Ber. 5), 2013.

Techel, F., Stucki, T., Margreth, S., Marty, C., and Winkler, K.: Schnee und Lawinen in den Schweizer Alpen. Hydrologisches Jahr 2013/14, WSL-Institut für Schnee- und Lawinenforschung SLF Davos: 87 pages (WSL Ber. 31), 2015a.

Techel, F., Zweifel, B., and Winkler, K.: Analysis of avalanche risk factors in backcountry terrain based on usage frequency and accident data in Switzerland, Nat. Hazards Earth Syst. Sci., 15, 1985–1997, doi:10.5194/nhess-15-1985-2015, 2015b.

Techel, F., Jarry, F., Kronthaler, G., Mitterer, S., Nairz, P., Pavšek, M., Valt, M., and Darms, G.: Avalanche fatalities in the European Alps: long-term trends and statistics, Geographica Helvetica, 71, 147–159, doi:10.5194/gh-71-147-2016, 2016.

Valt, M. and Cianfarra, P.: La stagione invernale 2013/2014. Innevamento e attivitá valanghiva sulle Alpi, Neve e Valanghe, 81, 10–19, 2014.

Wilks, D.: Statistical methods in the atmospheric sciences, vol. 100 of *International Geophysics Series*, Academic Press, San Diego CA, U.S.A, 3rd edn., 2011.

Williams, K.: Credibility of avalanche warnings, Journal of Glaciology, 26, 93–96, doi:10.1017/S0022143000010625, 1980.

Winkler, K. and Techel, F.: Users rating of the Swiss avalanche forecast, in: Proceedings ISSW 2014. International Snow Science Workshop, Banff, Canada, pp. 437–444, 2014.

Winkler, K., Fischer, A., and Techel, F.: Avalanche risk in winter backcountry touring: status and recent trends in Switzerland, in: Proceedings ISSW 2016. International Snow Science Workshop, Breckenridge, Co., pp. 270–276, 2016.

Zenkl, G.: 10 Jahre Lawinenwarndienst Niederösterreich, Saisonbericht der österrreichischen Lawinenwarndienste 2015/2016, pp. 192–193, 2016.