Backcountry Difficulty Property Dataset (BDPD)

User Manual

Günter Schmudlach

26. July 2020

(V1.2.2)

Abstract

The document introduces a training dataset, that powers modeling **difficulty degrees** of backcountry ski routes. The difficulty scale of the Swiss Alpine Club is based on four major criteria: Slope angle, exposure to fall down, terrain form and space conditions. The dataset contains the response variable (the difficulty degree) and a list of covariates for approximately 1200 backcountry ski tours of Switzerland. Furthermore the document adds some recommendations to be considered, if a model is trained from the dataset.

Content

1 Introduction

1.1 Difficulty degree

When mountaineers opt for a particular backcountry ski tour, the difficulty degree is an important information element. Its crucial that the skills of the mountaineers match the challenges that the route poses.

In Switzerland the rating is usually based on the Labande scale from the Swiss Alpine Club. It consists of 7 major degrees.

- L: leicht (easy)
- WS: wenig schwierig (little difficult)
- ZS: ziemlich schwierig (pretty difficult)
- S: schwierig (difficult)
- SS: sehr schwierig (very difficiult)
- AS: ausserordentlich schwierig (extremely difficult)
- EX: extrem schwierig (difficulty degree reserved for extreme skiing)

Most difficulty degrees can be refined by a plus (+) or minus (-) sign. Consequently 18 minor difficulty levels result:

diff	Level	Meaning	Steepest slope angles
1	L	leicht	under 30°
2	L+	leicht	
3	WS-	wenig schwierig	
4	WS	wenig schwierig	30...35°
5	WS+	wenig schwierig	
6	ZS-	ziemlich schwierig	
7	ZS	ziemlich schwierig	35...40°
8	ZS+	ziemlich schwierig	
9	S-	schwierig	
10	S	schwierig	40...45°
11	S+	schwierig	
12	SS-	sehr schwierig	
13	SS	sehr schwierig	45...50°
14	SS+	sehr schwierig	
15	AS-	ausserordentlich schwierig	
16	AS	ausserordentlich schwierig	50...55°
17	AS+	ausserordentlich schwierig	
18	EX	extrem schwierig	Over 55°

The degrees AS and EX have not a lot of practical meaning as they are reserved for extreme skiers. Currently the described dataset holds no route more difficult then level 13.

1.2 Criteria

The difficulty levels are usually attributed by “experts”. Often they are authors of mountain guide books and/or mountain guides. When an “expert” rates a route, he should apply four major criteria:

1. Slope angle: The most steep section of a ski tour gives a first idea about the difficulty degree.
2. Exposure to fall down: Exposure to down fall has an impact on the subjective perception of the difficulty.
3. Terrain form: Usually open, uniform and gentile slopes are considered easy. If the terrain holds obstacles, like gullies, sharp ridges or cliffs, its considered to be more difficult.
4. Space conditions: A particular section of a route can be narrow or wide. The more narrow (e.g. on ridges, in gullies), the more difficult the route.

As soon as the major difficulty degree is defined, the “expert” can apply three subordinated criteria to refine the degree with a plus (+) or minus (-) sign:

1. Difficult orientation during ascent and descent.
2. The course of the route is not well visible.
3. Route errors can hardly or not at all be corrected.

The difficulty degrees are approximate values in good snow and weather conditions.

1.3 Foot sections

The scale furthermore defines, that the difficulty degree only refers to the route sections that can be handled with skis. Sections that can only be negotiated on foot must be assessed using an other scale (the alpine technical scale).

In the Alps many backcountry ski tours end in a so called **ski depot**, where the skis are left. From the ski depot the remaining section to the top of the mountain must be handled on foot. On a few ski tours there can be foot sections long before the top is reached. So we can distinguish between two types of foot sections:

1. Mandatory foot sections: Such sections must be negotiated, if the mountaineer wants to accomplish the route.
2. Optional foot section: This is the final section from the ski depot to the top. We call it optional, because the skiing part of the route can be accomplished without the final foot section.

We assume the “experts” include mandatory sections and exclude optional sections, when they attribute a difficulty degree to a route.

1.4 Point Filtering

Terrain properties at particular point can be misleading on dams and on streets within harsh terrain. Therefore the dataset contains a filter file that indicates, whether a particular point has to be included/excluded from difficulty calculation:

- If a route point is closer then 20 m to a dam, its excluded.
- If a route point is closer then 20 m to a street and below 1600 m, its excluded.

Information about dams and streets comes from OpenStreetMap.

1.5 Discussion

Only one of four criteria (slope angle) is of quantitative nature, all other criteria are of qualitative nature. Rating difficulty degrees of backcountry ski routes is therefor a subjective process. The result depends on other factors:

- Snow and weather conditions met during the ascent.
- How far the “expert” wanted and could follow an optimal line during the ascent.
- Skills and mood of the “expert”.
- General “calibration” of the “expert”. “Experts” can differ considerably in the way they opt for a

difficulty degree. Some are known to rate more aggressively then others.

- Handling of foot sections: A section can be considered as foot section or as ski section. There are no objective criteria to distinguish such sections. Additionally there is no common understanding, if foot sections should be included or excluded from rating.

If difficulty degrees from different sources (like Guide books, Internet) are compared there can easily be differences up to 3-4 levels. Its important to underline that the difficulty degrees contained in the dataset are by no means consistent.

2 Point Properties

2.1 Motivation

A route consists of a number of points. The routes of the presented training dataset were resampled at a distance of 10 m. If the average route length is 6 km, it can be described by 600 equidistant points.

In chapter 1.2 we introduced four criteria to be applied for the difficulty degree attribution. The current chapter presents 10 properties linked to one or more of theses criteria. Each property is presented by a table structured as follows.

Name	**Abbreviation**	**Name**	
Description	Description of the property.		
Comment	A comment about the property.		
Values	Data type	Value range	No data value
Reference	Reference to more information about the property.		
Criteria	The criteria the property emulates		
Usage	0-3 Stars	Recommendations for the usage of the property.	
Copyrights	Copyrights of the raw data		

2.2 Slope Angle (SA)

Name	**SA**	**Slope Angle**	
Description	The slope angle derived from a DEM with 10 m resolution.		
Comment	The property and the criteria match 1:1		
Values	Decimal	0..90°	-9999
Reference	gdaldem (slope)		
Criteria	Slope angle		
Usage	***		
Copyrights	© Swisstopo		

2.3 Plan Curvature (PLANC)

Name	**PLANC**	**Plan Curvature**	
Description	The planar curvature calculated from a DEM with resolution 10 m.		
Comment	Negative values indicates convexity (n), positive values indicate concavity (u). Caution: In order to find an optimal scaling use GRASS and not ArcGIS to calculate PLANC.		
Values	Decimal	-100..100	-9999
Reference	r.param.scale(size=7, method=planc)		
Criteria	Terrain form, space conditions		
Usage	*	Priority should be given to FOLD.	
Copyrights	© Swisstopo		

2.4 Terrain Folds (FOLD)

Name	**FOLD**	**Terrain Folds**	
Description	Slope normal discontinuity raster. The raster shows folds (edges) in the terrain. Calculated from a DEM with 10 m resolution.		
Comment	Negative values indicates concavity (u), positive values indicate convexity (n).		
Values	Decimal	-180..180°	-9999
Reference			
Criteria	Terrain form, space conditions		
Usage	**		
Copyrights	© Skitourenguru		

The fold raster is calculated in 3 steps:

1. In a first step 10 **slope normals** are calculated on a circle with radius 10 m.
2. In a second step the angle between 5 pairs of **opposite slope normals** a calculated.
3. The maximal angle of all five angles gives the value of the fold raster.

The fold raster value is related to the MAXIC-Curvature r.param.scale(method=maxic).

2.5 Forest Density (FD)

Name	**FD**	**Forest Density**	
Description	Forest Density (in %) and a resolution of 20 m.		
Comment			
Values	Decimal	0..100%	-9999
Reference	Tree Cover Density (2015)		
Criteria	Space conditions		
Usage	*	Use with low priority.	
Copyrights	© ESA		

2.6 Maximal normal acceleration (FD_MAXNA)

Name	**FD_MAXNA**	**Maximal normal acceleration**	
Description	Maximal normal acceleration on a downfall trajectory.		
Comment			
Values	Decimal	0..400 m/s2	-9999
Reference	Avalanche terrain maps for backcountry skiing in Switzerland		
Criteria	Exposure to downfall		
Usage	*	Use with low priority, see FD_MAXV and FD_SUMV.	
Copyrights	© Skitourenguru		

The property is calculated through the following steps:

- A downfall trajectory of maximally 1 km length is calculated.
- An item of 75 kg falls down the downfall trajectory: Normal accelerations are recorded along the downfall trajectory. Normal accelerations cause injuries. Finally the maximal normal acceleration is extracted.

2.7 Sum of normal accelerations (FD_SUMNA)

Name	**FD_SUMNA**	**Sum of normal accelerationa**	
Description	Sum of normal accelerations on a downfall trajectory.		
Comment			
Values	Decimal	0..600 m/s2	-9999
Reference	Avalanche terrain maps for backcountry skiing in Switzerland		
Criteria	Exposure to downfall		
Usage	*	Use with low priority, see FD_MAXV and FD_SUMV.	
Copyrights	© Skitourenguru		

The property is calculated through the following steps:

- A downfall trajectory of maximally 1 km length is calculated.
- An item of 75 kg falls down the downfall trajectory: Normal accelerations are recorded along the downfall trajectory. Normal accelerations cause injuries. Finally the sum of normal accelerations is calculated.

2.8 Maximal Velocity (FD_MAXV)

Name	**FD_MAXV**	**Maximal Velocity**	
Description	Maximal velocity on a downfall trajectory.		
Comment			
Values	Decimal	0..80 m/s	-9999
Reference	Avalanche terrain maps for backcountry skiing in Switzerland		
Criteria	Exposure to downfall		
Usage	***		
Copyrights	© Skitourenguru		

The property is calculated through the following steps:

- A downfall trajectory of maximally 1 km length is calculated.
- An item of 75 kg falls down the downfall trajectory: Velocities are recorded along the downfall trajectory. Maximal velocity is extracted.

2.9 Sum of Velocities (FD_SUMV)

Name	**FD_SUMV**	**Sum of Velocities**	
Description	Sum of velocities on a downfall trajectory.		
Comment			
Values	Decimal	0..3000 m/s	-9999
Reference	Avalanche terrain maps for backcountry skiing in Switzerland		
Criteria	Exposure to downfall		
Usage	**		
Copyrights	© Skitourenguru		

The property is calculated through the following steps:

- A downfall trajectory of maximally 1 km length is calculated.
- An item of 75 kg falls down the downfall trajectory: Velocities are recorded along the downfall trajectory. Sum of velocities is calculated.

2.10 Corridor Width (CW)

Name	**CW**	**Corridor Width**	
Description	Width of a corridor modeled around the route. The property indicates, if a particular route section is narrow or width.		
Comment	CW is only available for routes in Switzerland. In all other regions of the Alps CW won’t be available for the years to come.		
Values	Decimal	0...500 m	-9999
Reference	Andreas Eisenhut: Skitourenplanung auf Knopfdruck		
Criteria	Space conditions		
Usage	**	Don’t use (see comment)	
Copyrights	© A. Eisenhut		

2.11 Risk (RISK)

Name	**RISK**	**Risk**	
Description	The product of SA (Slope angle) and FD_MAXV (maximal velocity on down fall trajectory)		
Comment	SA can be seen as a proxy for the probability to fall down. FD_MAXV can be seen as the consequences to fall down. The product gives us an indicator to the risk to fall down. Risk is here defined as the product of the “probability of an event” and the “consequences of the event”.		
Values	Decimal		-9999
Reference			
Criteria			
Usage	***		
Copyrights	© Swisstopo		

3 Covariates

3.1 Motivation

On the output side of the training dataset we have the difficulty level (**diff**) for each of the 1200 routes. On the input side we have a collection of route point properties. Before we can train a model, we need to combine the point properties along the route into some meaningful covariates valid for the route as a whole.

Usually the **most critical section** (e.g.: steep, exposed, narrow) serves as criteria for the difficulty degree of the route as a whole. However for two reasons its not advisable to just look for extreme values (min, max) along the course of the route:

1. Even if the route has been digitized with great care, it can be suboptimally placed. Consequently there can be extreme values that are meaningless for the most critical route section. An example: The most steep point of the dataset is 80.2° steep. Such slope angle is impossible and a typical runaway value.
2. A critical section (e.g.: steep, exposed or narrow) can be shorter or longer. Theoretically the difficulty doesn’t increase with rising length of the critical section. Practically a long critical section will be perceived by mountaineers differently then a short critical section. Consequently the length of the critical section must be included when meaningful covariates for the route are derived.

As mentioned in chapter 1.3 optional foot sections shouldn’t be taken into account when calculating covariates. The dataset contains a file (Switzerland_EndPoints.csv) that indicates for each route the point that serves as ski depot. Each ski depot has a type. The type can be:

1. Type=1: There is no ski depot. Its possible to climb the top with skis.
2. Type=2: There is a ski depot. The value **count** points to the ski depot.
3. Type=3: Technically there is a ski depot, but its so close to the top, that we can suppose its possible to climb to the top.

As mentioned in chapter 1.4 some points must be included resp. excluded from difficulty calculation. The file Switzerland_RouteDifficultyFilter.csv contains the according information (attribute ok).

The dataset contains four series of covariates. These series follow a different strategy to combine **magnitude** and **length** of critical sections. In the next chapters all series are presented.

3.2 Maximum

The average property of the most critical n points define the covariate. The value n is 1, 2, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100.

covariate = mean(max$_n$ ($|property|$))

An example with slope angle: If n=10, the mean slope angle over the 100 steepest meters of the route define the covariate.

The following diagram shows the distribution of the mean slope angle over the steepest 100 meters for 12 difficulty levels.

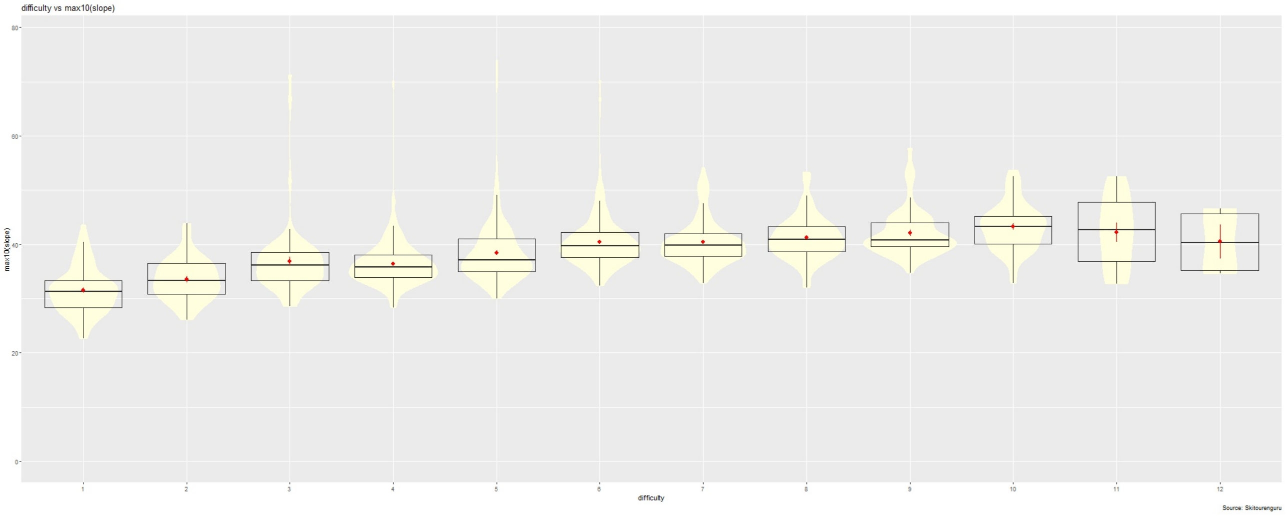


Fig. 1: Distribution of slope angles (y) over 12 difficulty levels (x)

3.3 Weighted Maximum

The average property of the most critical n points define the covariate. The value n is 10, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200. In contrast to the covariate of chapter 3.2 the values are weighted. Weights fall linearly from 1 to 1/n The most critical property is weighted by 1, the least critical value is weighted by 1/n.

covariate = weighted.mean(max$_n$ ($|property|$), weights)

weights = seq(from=1, by=-1/n, length=n)

An example with slope angle: If n=10, the mean slope angle over the 100 steepest meters of the route define the covariate. The most critical value receives is weighted by 1, the least critical value is weighted by 0.1.

In contrast to the covariates defined under 3.2 this covariate handles consistently the transition area between point n and point n+1.

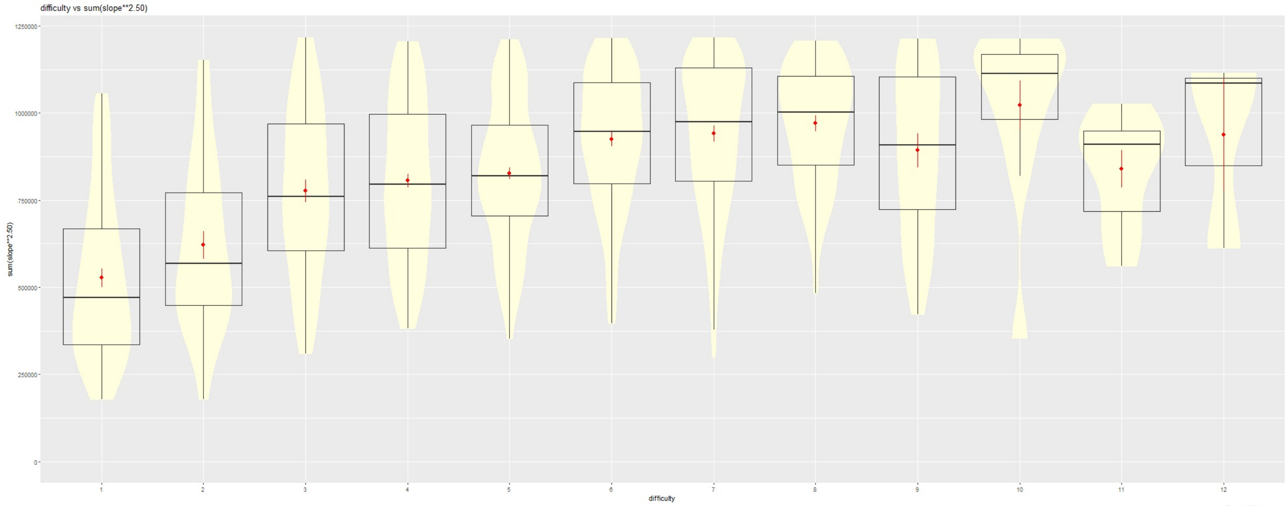
3.4 Exponential

The sum of the exponentiated properties define the covariate. The exponent e is 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5 or 5.

covariate = $\sum |property|^e$

If e=1 the covariate corresponds to the area under the curve. If e>1 higher values are weighted over-proportionally then lower values.

The following diagram shows the distribution of the summed exponentiated (e=2.5) slope angles for 12 difficulty levels:



Source: Skitourenguru

Fig. 2: Distribution of slope angles (y) over 12 difficulty levels (x)

3.5 Time series

This bundle of covariates is calculated by a “time serie feature extraction” library:

https://github.com/nelsonroque/tsfeaturex

In total 69 covariates per property are calculated.

4 Training dataset

The training dataset consists of the following files:

1. ReponseVariable/Switzerland.csv: Holds for each route (defined by its **id**) the response variable (the difficulty level **diff**).
2. Covariate_Max/{*Property*}_n.csv: Holds for each route (defined by its **id**) the covariate called “maximum” (see chapter 3.2) for the according property. The parameter n is the number of points taken into account. Close to the *.csv file there is a *.jpg file that shows the distribution of the covariate.
3. Covariate_Wmax/{*Property*}_n.csv: Holds for each route (defined by its **id**) the covariate called “weighted maximum” (see chapter 3.3) for the according property. The parameter n is the number of points taken into account.
4. Covariate_Exp/{*Property*}_e.csv: Holds for each route (defined by its **id**) the covariate called “exponential” (see chapter 3.4) for the according property. The parameter e is the exponent. Close to the *.csv file there is a *.jpg file that shows the distribution of the covariate.
5. Covariate_TS/{*Property*}_e.csv: Holds for each route (defined by its **id**) the covariate called “time serie” (see chapter 3.5) for the according property.

The training dataset holds as well the raw properties at the route points:

1. Raw/Switzerland_{*Property*}.csv: Property values on equidistant points along all routes. Points belonging to the same route have the same **id**. The point sequence in the file corresponds to the sequence during the ascent.
2. Raw/Switzerland_EndePoints.csv: A list of **count** values, that point for each route (defined by its **id**) to the ski depot.
3. Raw/Switzerland_RouteDifficultyFilter.csv: A list of bool values (attribute: ok), indicating, if the according point has to be included (1) or excluded (0) from difficulty calculation.

The raw data are not needed to train a model. However the raw data could be used to derive new covariates.

5 Recommendations

When a model is trained from the training dataset its important to keep in mind the subjectivity of the difficulty degree (see chapter 1.5). It is pointless to develop a refined model that captures all extravagances of the training dataset. We must be aware that many of the routes are not well rated or even wrongly rated. Instead of optimizing a model criterion we should rather focus on the basic principles applied by the “experts”. The model should **postulate a standard**, that defines how “experts” do rate and should rate the difficulty degrees. Such a model could be called “normative”. Of course the standard must take into account the general patterns contained in the training dataset.

Consequently that means that the model should be simple, transparent and highly generalized. That means in particular:

- Select only a few covariates (may be only 3 to 4).
- Covariate selection should be based on an objective criterion.
- Try to avoid the covariate CW. CW is not available outside of Switzerland. A model that can be used only in Switzerland is pointless.
- It must be possible to communicate the inside of the model. Use a white box model, avoid black box models.
- In a first step use a simple model. An obvious example is linear regression of order 1. In a second step develop eventually a more complex model (eventually a GAM).
- Make sure you rather under- then overfit the training data.

The model to be developed defines formally the way “experts” **do rate** and **should rate** the difficulty degrees of backcountry ski tours.

The same dataset was used by the bachelor thesis of Brunner & Bittel (2020). They applied linear regression analysis and a random forest algorithme.

6 Further Reading

Swiss Alpine Club: SAC Schwierigkeitsskala für Skitouren

Bunner F., Bittel M.: Künstliche Intelligenz & Skitouren - Automatische Bestimmung des Schwierigkeitsgrades von Skitouren; Diplomarbeit an der Fachhochschule Nordwestschweiz; Brugg, 2020