Avalanche Risk Property Dataset (ARPD)

User Manual

Günter Schmudlach

19. October 2020

(V3.0.12)

Abstract

The document introduces a generic dataset, that powers avalanche risk calculations. The dataset consists of two OGR compliant point vector datasets. The first dataset contains information about avalanche accident points (failure data). The second dataset contains information about travel points of the back-country skier community (success data). Each point of both datasets provides a geographic location and a number of properties. Special emphasis is given to the description of the point properties. Both datasets refer to Switzerland. The purpose of the data set is to derive knowledge about avalanche risk.

Content

1 Introduction

Where and when do avalanche accidents in a backcountry skier context occur? In a classical approach scientists search for specific patterns within the property data of avalanche accidents. However accidents in a backcountry skier context only occur, if there is backcountry skier traffic. Therefore its mandatory to relate to knowledge about accidents to the knowledge about the underlying backcountry skier traffic.

The **Avalanche Risk Property Dataset (ARPD)** consists of two datasets:

1. The first dataset describes 1037 avalanche accidents of Switzerland. In all these accidents humans were involved All accidents occurred in a backcountry skier context. Accidents that occurred in a freerider context were discarded. In 95 % of the cases a winter sportsman triggered the avalanche. The data cover all severe accidents from the winter 2001/2002 to the winter 2018/2019. Approximately 17 % of the accidents had fatal consequences. Each accident is described by 7 points. The first point refers to the highest point of the release area. The remaining points describe a downhill trajectory starting at the first point. Points are located in a 10 m distance. The downhill trajectory describes a line covering the most likely **release area** of the avalanche.
2. The second dataset describe approximatively 5 million transition points of backcountry skiers in Switzerland. The data come from GPS track recordings of effectively undertaken skitours. The GPS tracks were uploaded from contributors to the platforms skitourenguru.ch, gipfelbuch.ch and camp2camp.org. A complex filtering process makes sure the dataset only contains information about backcountry skitours. Points were resampled with a constant distance of 10 m along the routes.

The first dataset provides information about **failure points**. Failure point means an accident occurred at this point. The second dataset provides information about **success points**. Success point means that the point could be passed without triggering an avalanche. For each point of both datasets a number of properties are provided. By comparing the properties of failure points to the properties of success points its possible to deduce knowledge about the **relative avalanche risk**.

ISO norms define risk as the **effect** of **uncertainties** to **objectives**. We define two objectives:

1. To prevent backcountry skiers to trigger an avalanche resp. to be caught by an avalanche.
2. To keep a reasonable space of freedom of movement to backcountry skiers.

The knowledge can be used to model tools that direct the backcountry skier community in time and space to less riskier domains.

2 Property Description Template

The following chapters will present all properties of accident and transition points. An introducing table with the following elements gives an overview to the property:

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.		
Redundancy	Information about redundancy to other properties.		
Usage	0-3 Stars	Recommendations for the usage of the property.	
Copyrights	Copyrights of the raw data		

The section **Usage** indicates a number of stars. The number of stars depend on two criteria:

1. The predictive value of the property: A property is a good **explanatory variable** for the risk, if failure data and success data show a clear trend and if failure data and success data are fundamentally different.
2. Availability: Availability of the data throughout the Alps. The availability of data includes as well copyrights limitations.

After the table a longer description about the property is given if required. Its followed by a histogram of the property: **Raw data** are displayed with **dashed lines**, **smoothed data** are displayed with a **solid line**. The following colors are used:

1. **Blue**: A histogram of the properties at the terrain usage points (success data). The **dashed line** shows raw data, the **solid line** shows smoothed data.
2. **Yellow**: A histogram of the properties at the accident points (failure data). The **dashed line** shows raw data, the **solid line** shows smoothed data.
3. **Red**: A histogram of the quotient from the properties at the accident points to the properties at the transition points. The **dashed line** shows raw data, the **solid line** shows smoothed data. The **dotted line** comes from the division of the **smoothed accident line** and **smoothed terrain usage line**.

The vertical axis shows the **mean normalized frequency**. The mean normalized frequencies is calculated by dividing the **initial frequency** by the **mean frequency**.

Smoothing is done with Kernel Density Estimation (KDE). KDE needs the definition of a bandwidth. The bandwidth is normally 10% of the range covered by the horizontal axis. For AOF, DI, IDI and RDL its 50%.

The histogram provides a first insight to the distribution of the property values. It will be followed by a preliminary interpretation.

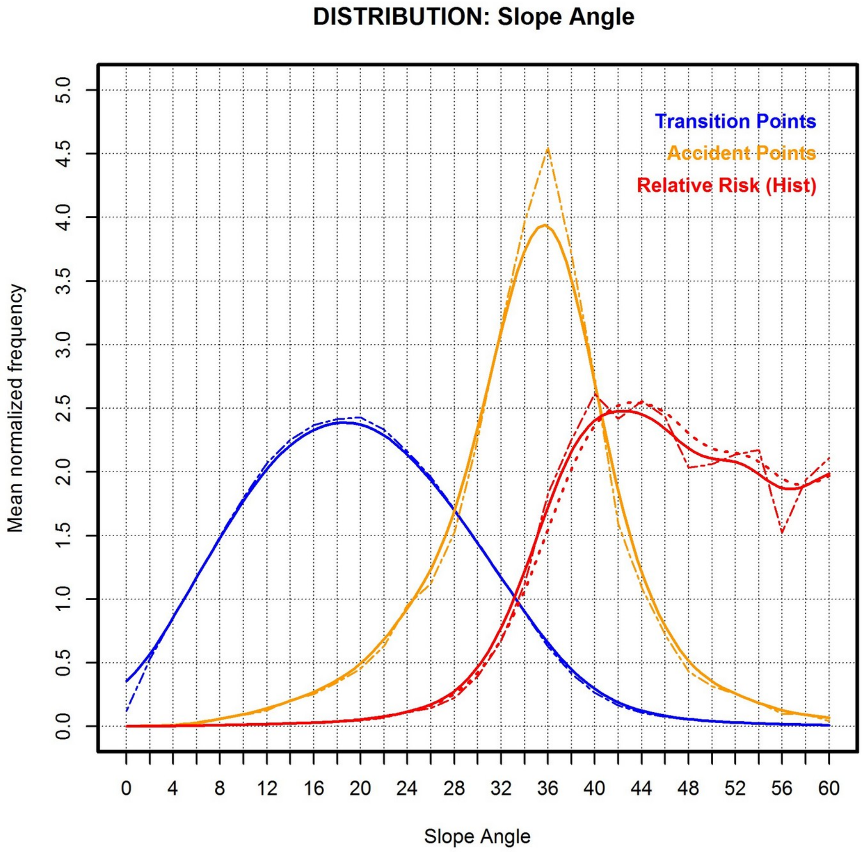
3 Data Bias

Knowledge about an eventual data bias is collected on Skitourenguru.

4 DataTerrain (Avalanche probability)

4.1 Slope Angle (SA)

Name	SA	Slope Angle	
Description	The slope angle derived from a DEM with 10 m resolution.		
Comment			
Values	Decimal	0..90°	-9999
Reference	gdaldem (slope)		
Redundancy	TI, MSA*, HP, SP		
Usage	***	Priority should be given to TI.	
Copyrights	© Swisstopo		

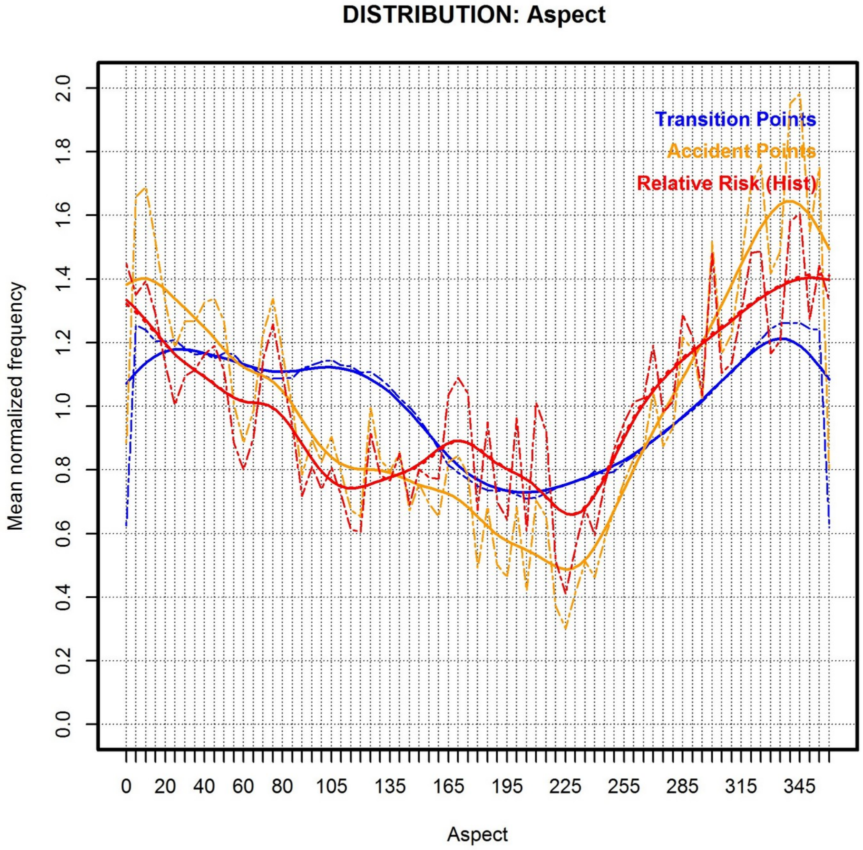


Interpretation

From former accident data analysis it was already known, that most accidents occur around 36-40°. If accident data is related to terrain usage data, the curve moves to the right (see red curve). The highest relative risk is reached at 44°. Above 44° the relative risk slowly decreases. Uncertainties become high above 50°.

4.2 Aspect (ASPECT)

Name	**ASPECT**	**Aspect**	
Description	Aspect at the point.		
Comment	The aspect derived from a DEM with 10 m resolution.		
Values	Decimal	0..360°	-9999
Reference	gdaldem (aspect)		
Redundancy			
Usage	**	Priority should be given to AOF.	
Copyrights	© Swisstopo		

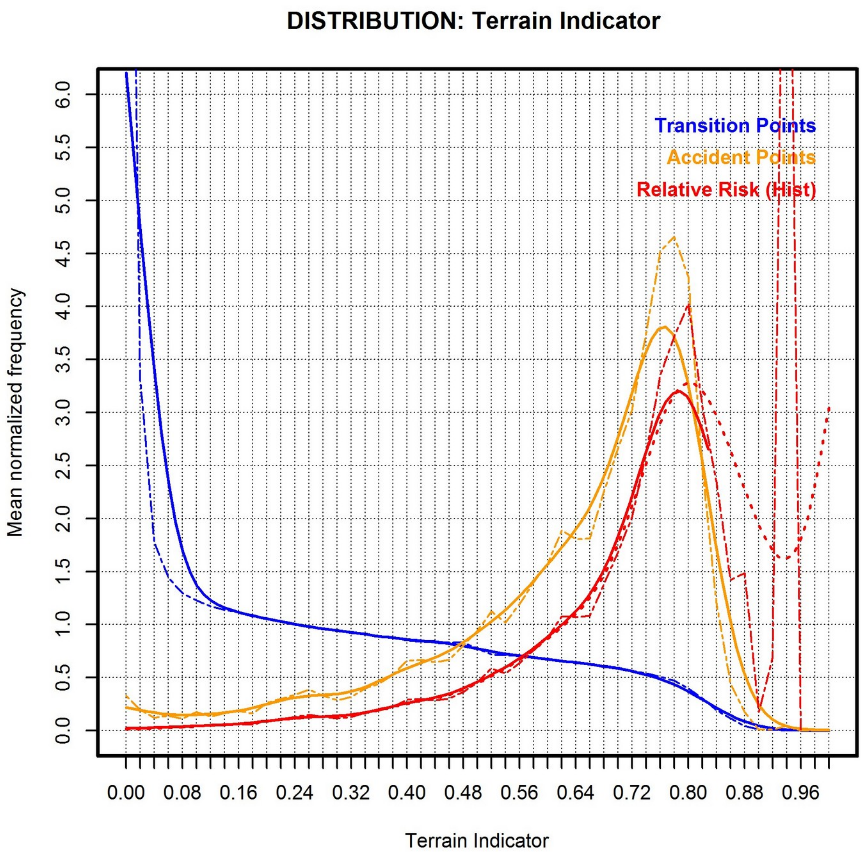


Interpretation

The lowest relative risks (0.7) can be found on SW-Slopes (225°). Highest relative risks (1.25) can be found on northern slopes That's 80% more then the minimal value. The difference is marked, but less then what was suggested in the past.

4.3 Terrain Indicator (TI)

Name	**TI**	**Terrain Indicator**	
Description	TI indicates how suitable a terrain point is to trigger an avalanche. MRSAR=100 m.		
Comment	TI is an immediate function of HP and SP. TI doesn't include directly the consequences of an avalanche (PBD, FD_*).		
Values	Decimal	0..1	-9999
Reference	Method for an Automatized Avalanche Terrain Classification		
Redundancy	ATH, SA, MSA*, HP, SP		
Usage	***	Important.	
Copyrights	© Skitourenguru		



Interpretation

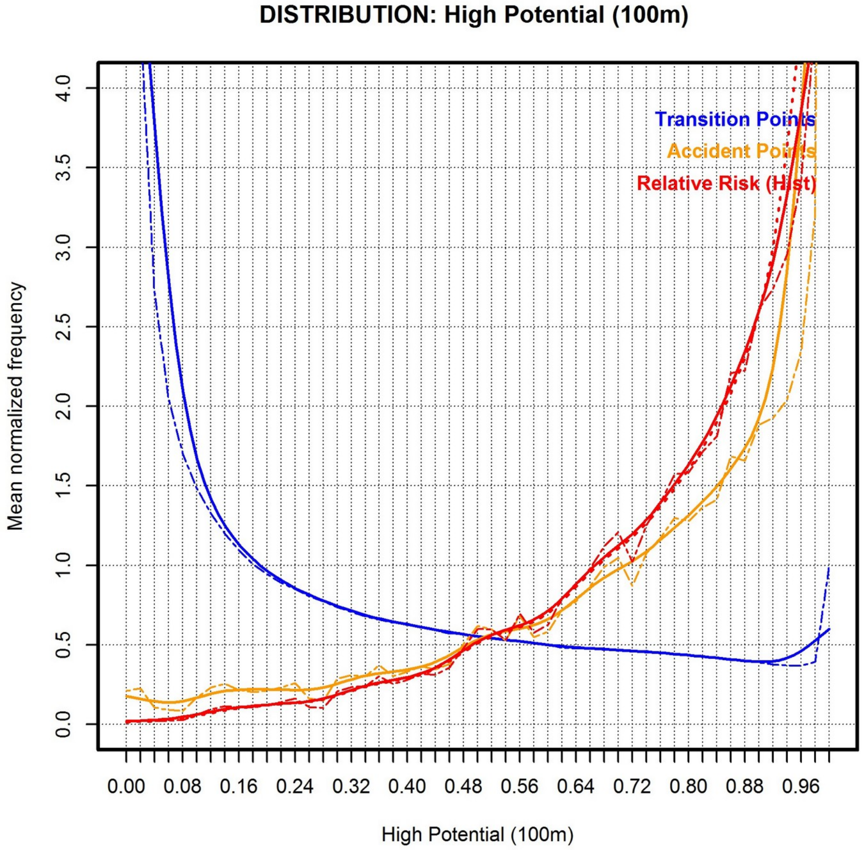
Failure data and success data follow a fundamentally different trend. The relative risk shows a tremendous rise in risk with rising TI. Above TI=0.8 the relative risk becomes uncertain.

4.4 High Potential (HP, HP70, HP150)

Name	HP	High Potential	
Description	HP gives a measure for the **avalanche potential** of the most dangerous spot within the slope the current point is member of.		
Comment	HP is an immediate precursor of TI. HP was calculated with different MRSAR (Maximal Relevant Slope Area Radius): • HP: 100 m • HP70: 70 m • HP150: 150 m		
Values	Decimal	0..1	-9999
Reference	Method for an Automatized Avalanche Terrain Classification		
Redundancy	SP, TI, SA, MSA*		
Usage	**	Priority should be given to TI.	
Copyrights	© Skitourenguru		

Each point is member of a slope called RSA (Relevant Slope Angle). All slope angles on the RSA are first converted to a potential. Therefore the knowledge of the risk distribution of SA is applied. HP is given then by applying the following formula:

HP = norm(mean(potentials) + 1.25 * sigma(potentials) where potential = risk(slopeAngle)



Interpretation

The relative risk shows a tremendous rise in risk with rising HP.

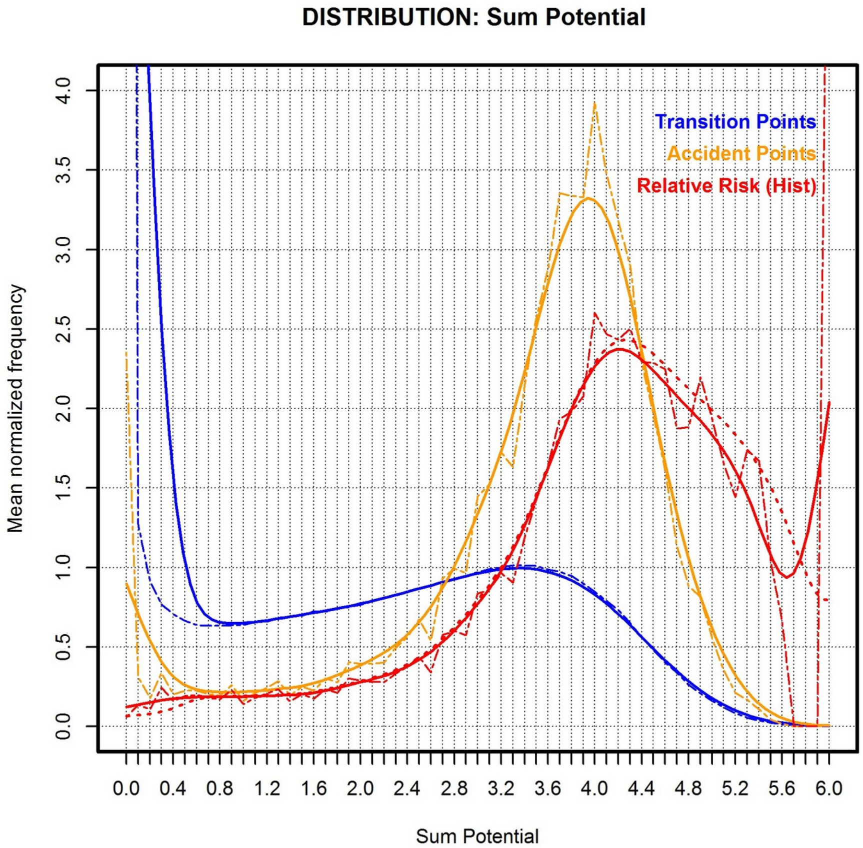
4.5 Sum Potential (SP)

Name	SP	Sum Potential	
Description	SP gives a measure for the **avalanche potential** of the whole slope the current point is member of. MRSAR=100 m.		
Comment	SP is an immediate precursor of TI.		
Values	Decimal	0..8	-9999
Reference	Method for an Automatized Avalanche Terrain Classification		
Redundancy	HP, SA, MSA*, PBD		
Usage	**	Priority should be given to TI.	
Copyrights	© Skitourenguru		

Each point is member of a slope called RSA (Relevant Slope Angle). All slope angles on the RSA are first converted to a potential. Therefore the knowledge of the risk distribution of SA is applied. SP is given then by applying the following formula:

SP = log(sum(potentials)) where potential = f(slopeAngle)

SP becomes a measure for the **size of the slope** or more exactly the slope's avalanche potential. As SP represents the size of the slope, its correlated to PBD (Projected Burial Depth). The larger the slope the higher are potential burial depths.

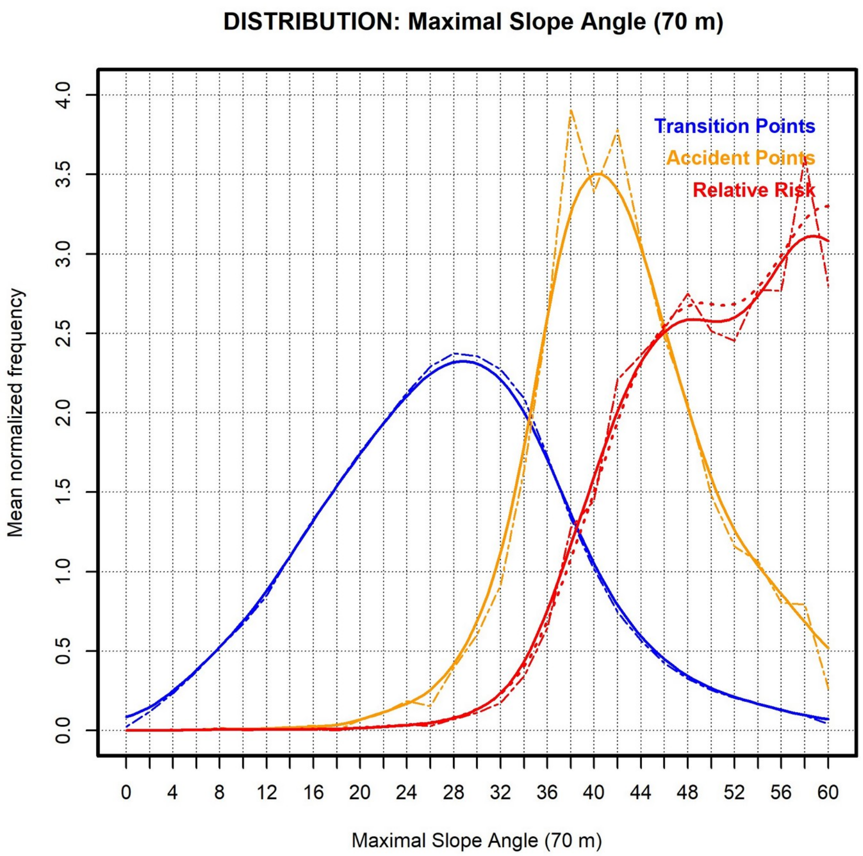


Interpretation

Failure data and success data follow a different trend. The relative risk shows a tremendous rise in risk with rising HP.

4.6 Maximal Slope Angle (MSA40, MSA70, MSA100, MSA150)

Name	**MSA**	**Maximal Slope Angle**	
Description	MSA indicates the **slope angle** at the most dangerous spot within the slope the current point is member of.		
Comment	MSA was calculated with different MRSAR (Maximal Relevant Slope Area Radius): • MSA40: 40 m • MSA100: 70 m • MSA100: 100 m • MSA150: 150 m The most dangerous spot is defined by the average slope angle of the 15% most dangerous raster cells. To identify the most dangerous raster cell “ceiling” is applied. Example: If we have 28 Pixels, the Math.Ceiling(28*0.15) = 5 most dangerous pixels would define the MSA.		
Values	Decimal	0..90°	-9999
Reference	[Method for an Automatized Avalanche Terrain Classification]()		
Redundancy	SP, HP, TI, SA		
Usage	**	Priority should be given to TI.	
Copyrights	© Skitourenguru		



Interpretation

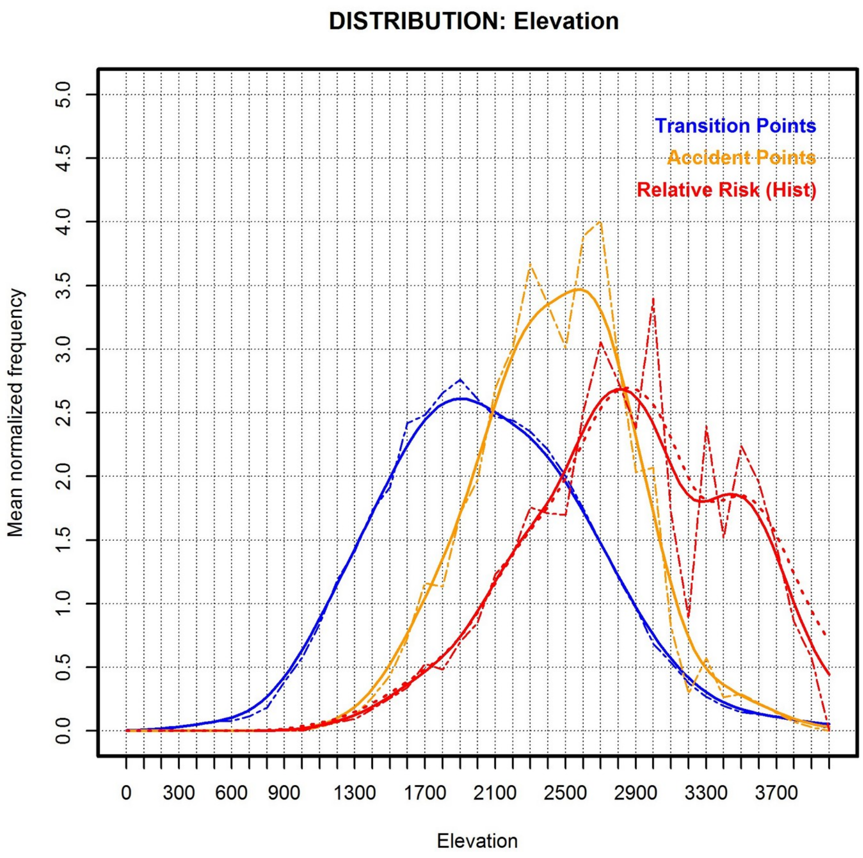
The relative risk shows a tremendous rise in risk with rising MSA*.

4.7 Minimal and Maximal Slope Aspect (MIN_ASPECT, MAX_ASPECT)

Name	**MIN_ASPECT, MAX_ASPECT**	**Minimal and Maximal Slope Aspect**	
Description	Indicates the interval of aspects contained in the slope assigned to the point. The interval starts at MIN_ASPECT goes in clockwise direction and ends at MAX_ASPECT.		
Comment	The assigned slope was calculated with a MRSAR (Maximal Relevant Slope Area Radius) of 100 m. In order to calculate the aspect interval only pixels with slope angle >25° are taken into account. Be careful when processing the interval: • For flat spots the values are set to NO_DATA • These are circular values: 360°=0°, that means MIN_ASPECT can be higher then MAX_ASPECT.		
Values	Decimal	0..360°	-9999
Reference	Method for an Automatized Avalanche Terrain Classification		
Redundancy	ASPECT		
Usage	**		
Copyrights	© Skitourenguru		

4.8 Elevation (ELE)

Name	**ELE**	**Elevation**	
Description	Elevation according to the DEM with 10 m resolution.		
Comment			
Values	Decimal	0..5000 m	-9999
Reference	swissALTI3D-10m		
Redundancy	None		
Usage	***	An important property.	
Copyrights	© Swisstopo		

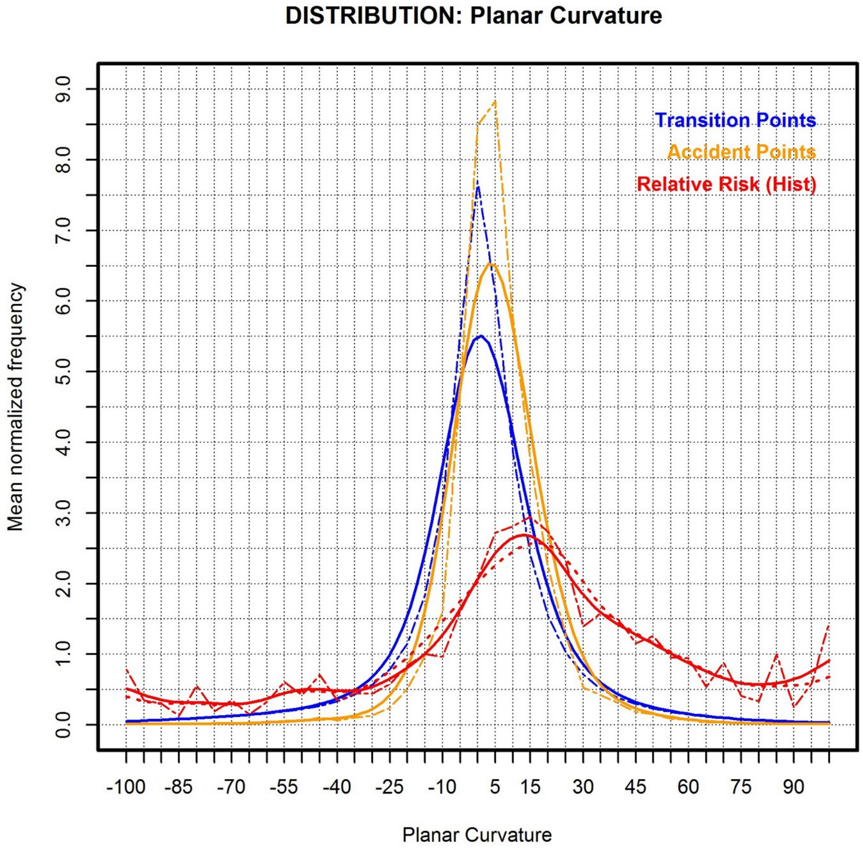


Interpretation

Failure data and success data follow a fundamentally different trend. The relative risk shows a tremendous rise in risk until 2700 m. From 2700 m on the further course of the relative risk becomes uncertain. Elevation is eventually the single most neglected property in past statistical avalanche assessment.

4.9 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)		
Redundancy	FOLD, TR		
Usage	**	Priority should be given to FOLD.	
Copyrights	© Swisstopo		



Interpretation

The data indicate higher risk for concavity (u) then for convexity (n), which applies to intuition. The success data and failure data follow a rather similar trend. Nevertheless the rise in risk is relatively prominent (from 0.5 to 2.5).

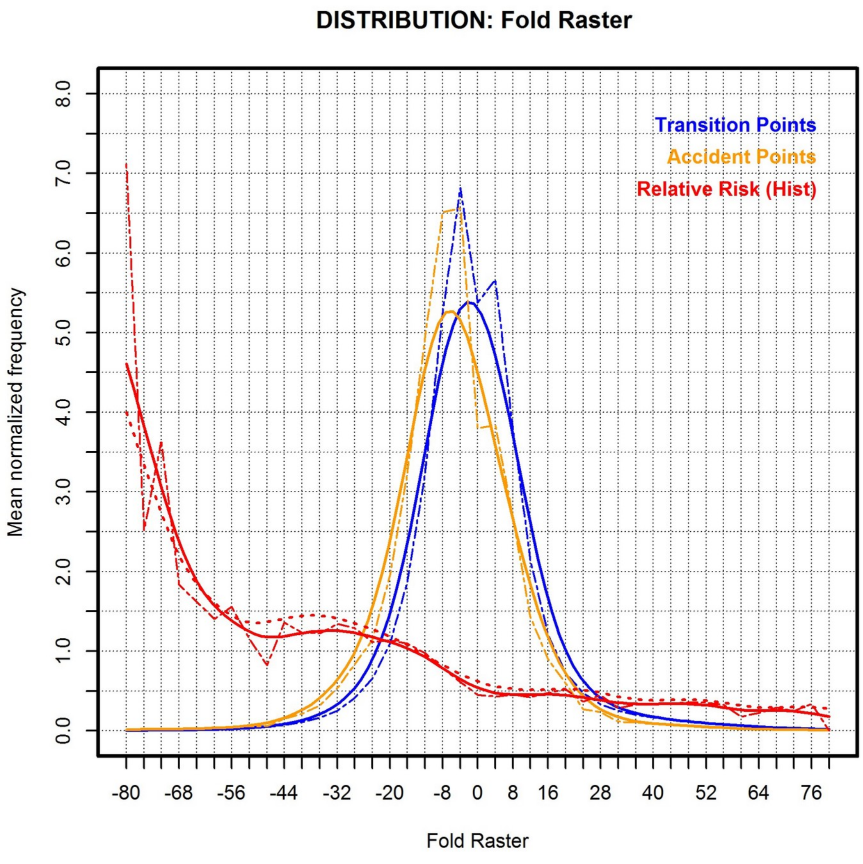
4.10 Terrain Folds (FOLD)

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			
Redundancy	PLANC, TR		
Usage	**	If used, give priority to this property over PLANC.	
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).

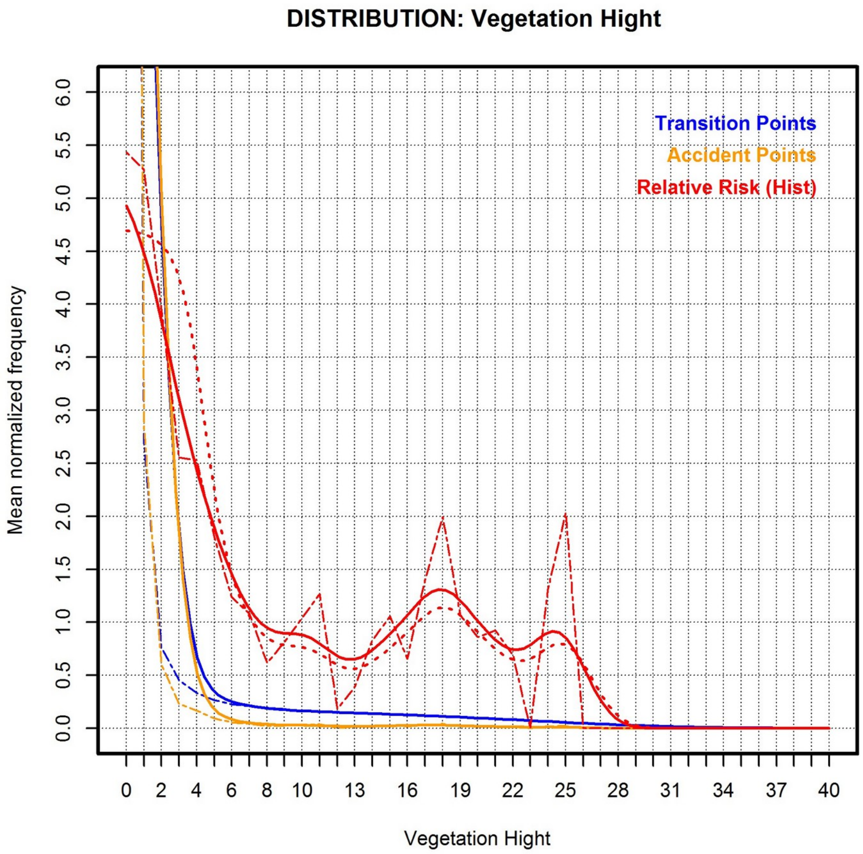


Interpretation

The data indicate higher risk for concavity (u) then for convexity (n), which applies to intuition. The success data and failure data follow a rather similar trend. Nevertheless the decline in risk is relatively prominent (from 3.5 to 0.25).

4.11 Vegetation Height (VH)

Name	**VH**	**Vegetation Height**	
Description	Vegetation Height Model and a resolution of 1 m.		
Comment	Data is only available for Switzerland.		
Values	Decimal	0..50 m	-9999
Reference	Vegetation Height Model (NFI)		
Redundancy	VD		
Usage	-	Don't use the dataset, its only available for Switzerland.	
Copyrights	© BAFU		

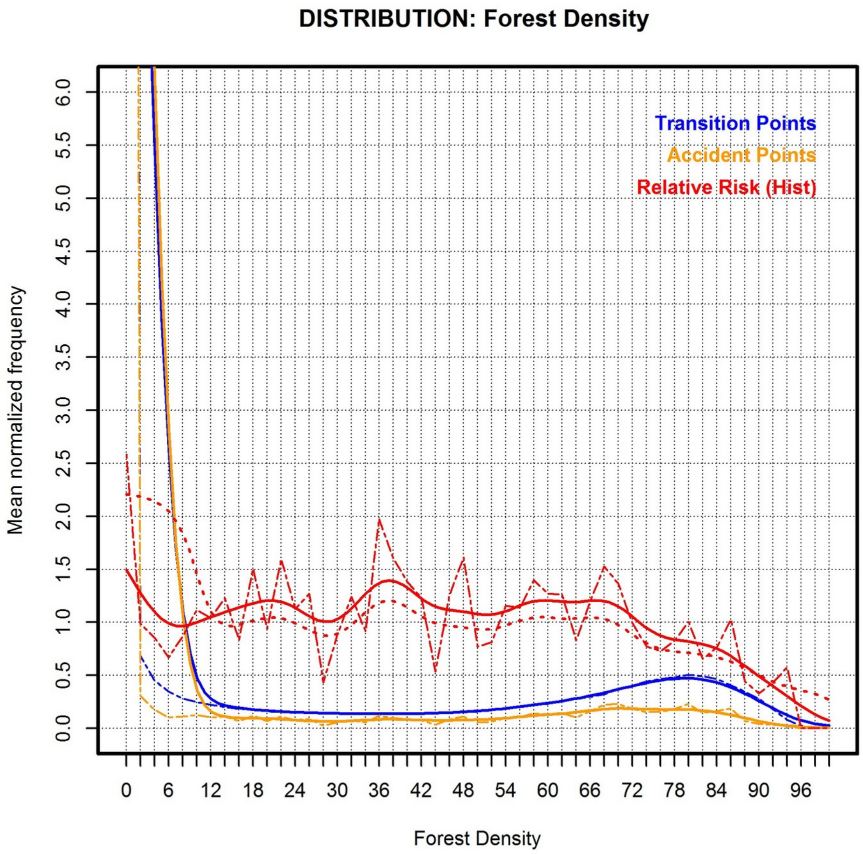


Interpretation

Between 0 and 10 m the protective value of forest constantly rises. From 10 m on there is no clear trend.

4.12 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)		
Redundancy	VH		
Usage	**	Use with medium priority.	
Copyrights	© ESA		

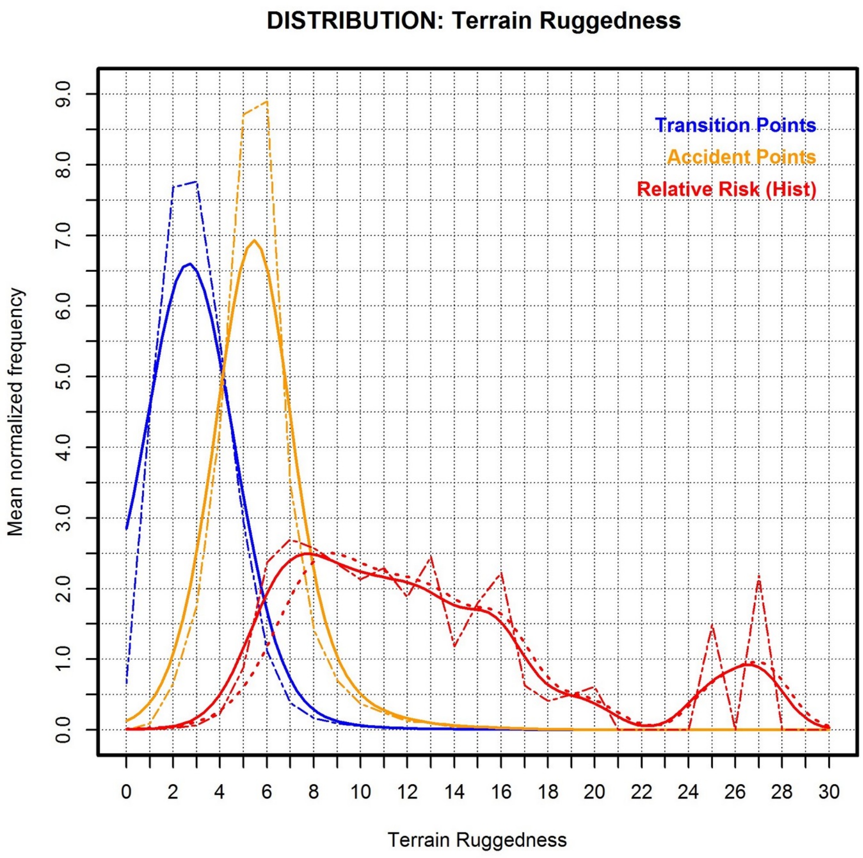


Interpretation

Forest starts to protect above vegetation density 72. There is a decline in risk from 72 to 90 (factor 2).

4.13 Terrain Ruggedness (TR)

Name	**TR**	**Terrain Ruggedness**	
Description	Terrain Ruggedness Index.		
Comment	TR is defined as the mean difference between a central pixel and its surrounding cells. TR is highly correlated to SA.		
Values	Decimal	0..30	-9999
Reference	gdaldem (TRI)		
Redundancy	FOLD, PLANC		
Usage	*	Use with caution.	
Copyrights	© Swisstopo		

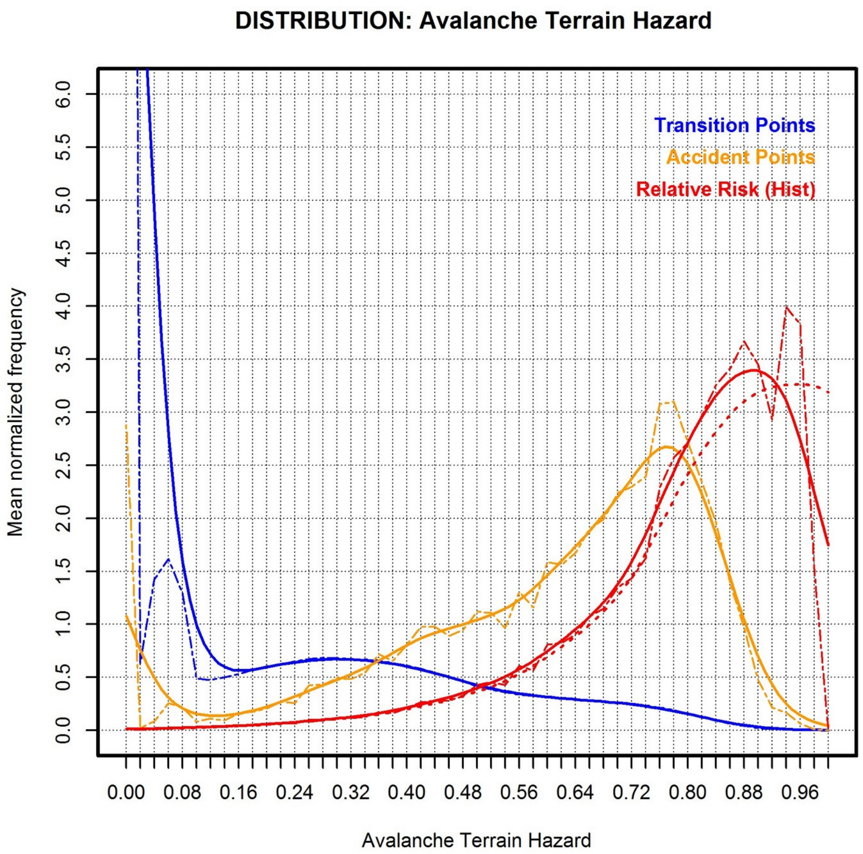


Interpretation

There is a difference between success data and failure data. We would expect a constant decline of risk with rising TR. However that's only the case above the value 8.

4.14 Avalanche Terrain Hazard (ATH)

Name	**ATH**	**Avalanche Terrain Hazard**	
Description	Terrain Classification of the SLF.		
Comment	ATH includes consequences of an avalanche (PBD, FD_*).		
Values	Decimal	0..1	-9999
Reference	[Avalanche terrain maps for backcountry skiing in Switzerland]		
Redundancy	TI, SA, HP, SP		
Usage	-	Don’t use for two reasons: 1. Data is only available for Switzerland. 2. License issue, as ATH is based on swissALTI3D with 5 m resolution. 3. Based on V1 of ATH.	
Copyrights	© SLF		



Interpretation

Failure data and success data follow a fundamentally different trend. The relative risk shows a tremendous rise in risk with rising ATH. Above ATH=0.9 the relative risk becomes uncertain.

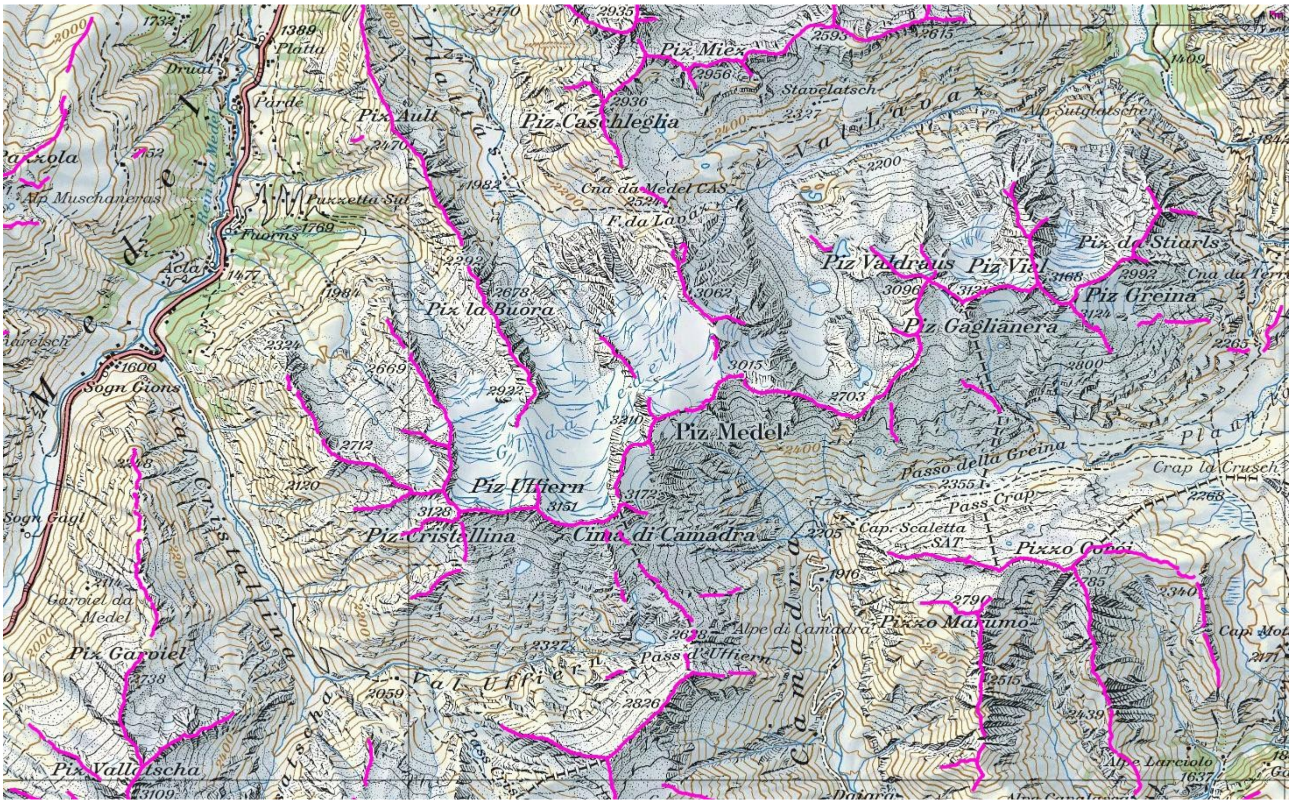
4.15 Distance to Ridge (DIST_RIDGE)

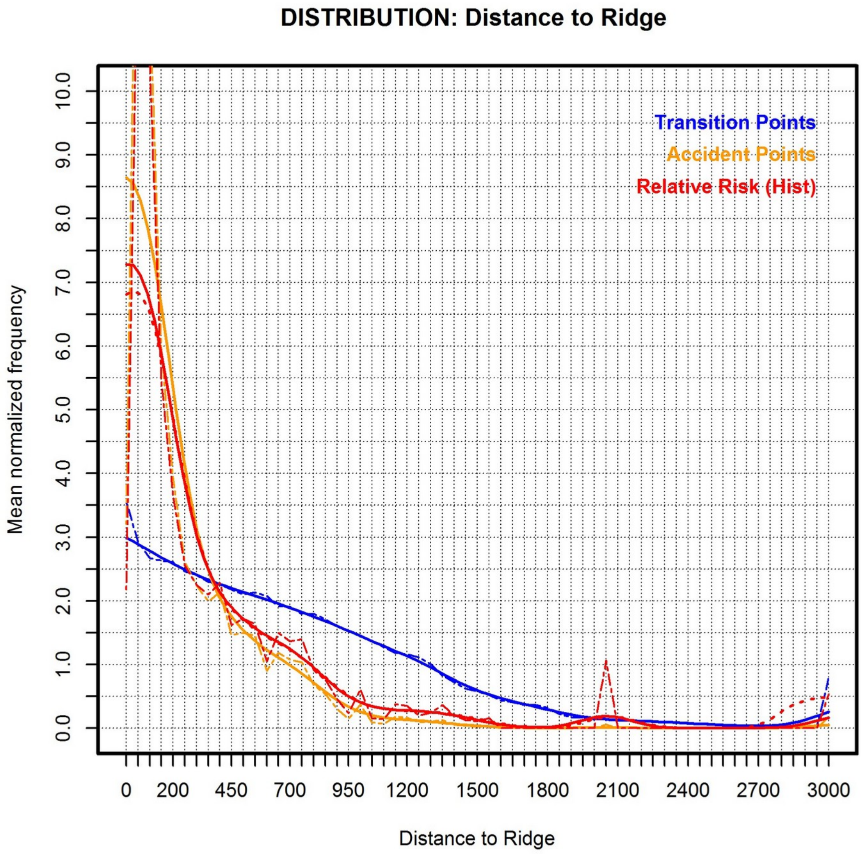
Name	**DIST_RIDGE**	**Distance to Ridge**	
Description	Distance to the next ridge, calculated from the DEM with 10 m resolution.		
Comment	All point with a distance larger then 3000 m will have the value 3000 m.		
Values	Decimal	0..3000	3000
Reference			
Redundancy	ELE		
Usage	***	Important property.	
Copyrights	© Skitourenguru		

Ridges are calculated with GRASS, following an approach that combines four different parameters:

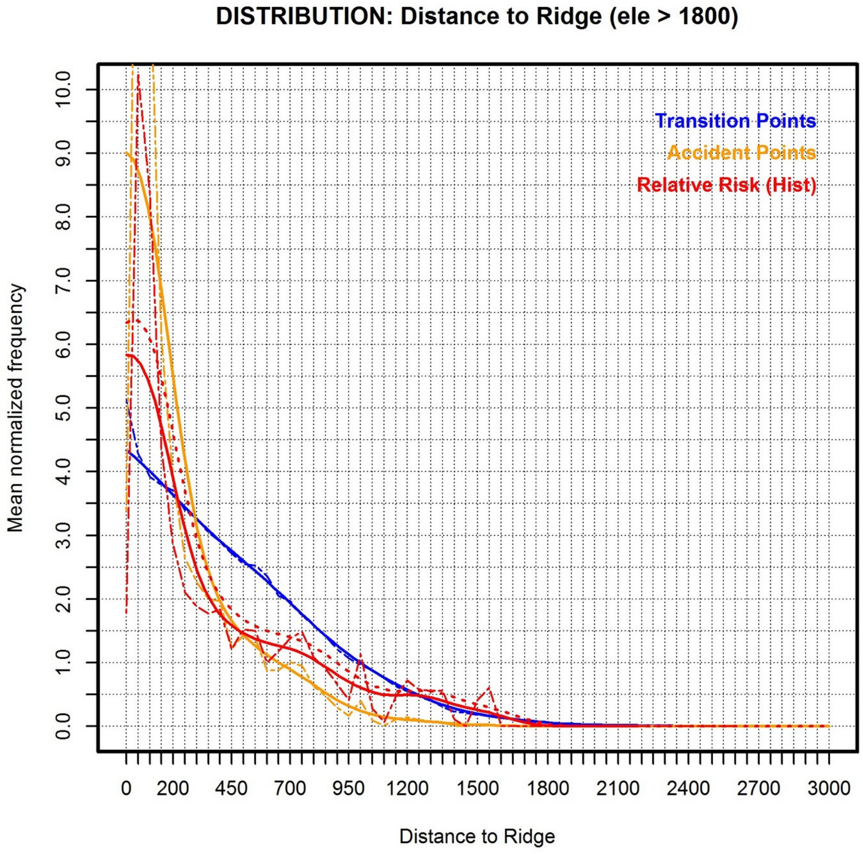
1. Water accumulation of the inverted DEM (r.watershed)
2. Absolute elevation
3. MAXIC-Curvature (r.param.scale(method=maxic))
4. Topex (r.horizon)

The normalized parameters are multiplied. The result is then converted with a threshold to a binary raster. In the last step the binary raster is vectorized. The following image gives an example of the result:





Points with low elevation have always a high distance to ridges. Therefore DIST_RIDGE can be correlated to ELE. The following graphic shows the histogram for points above 1800 m.



Interpretation

Even if we select only points above 1800 m the risk significantly decreases with distance to ridge.

4.16 Treeline (TL)

Name	**TL**	**Treeline**	
Description	The theoretical treeline elevation at the point		
Comment	Outside mountain areas the value is on 1800 m		
Values	Decimal	1800..2350 m	NA
Reference	GIS-analysis of tree-line elevation in the Swiss Alps suggests no exposure effect		
Redundancy			
Usage	*		
Copyrights	© WSL		

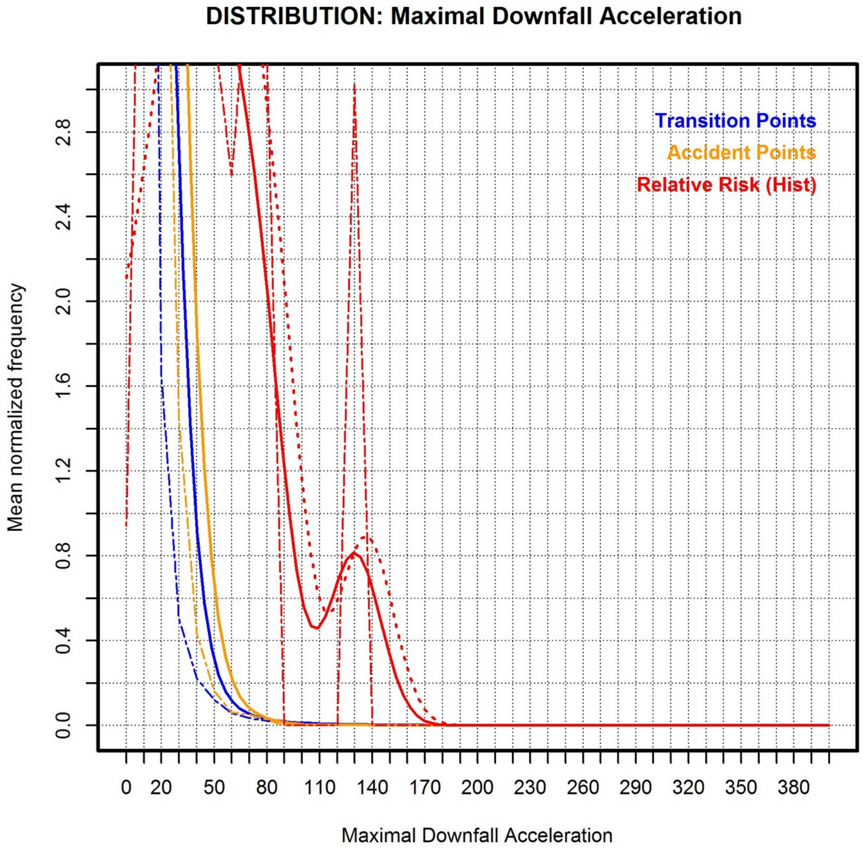
5 DataTerrain (Avalanche consequences)

5.1 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		
Redundancy	FD_*		
Usage	-	Don't use, see other properties FD_*.	
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.



Interpretation

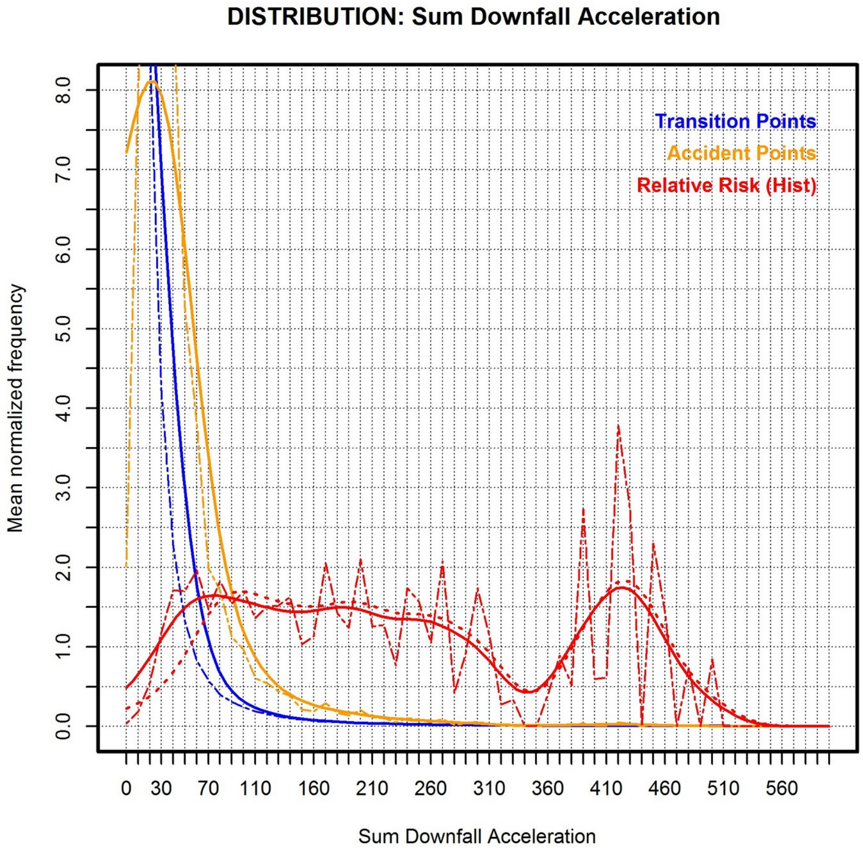
No clear trend.

5.2 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		
Redundancy	FD_*		
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.



Interpretation

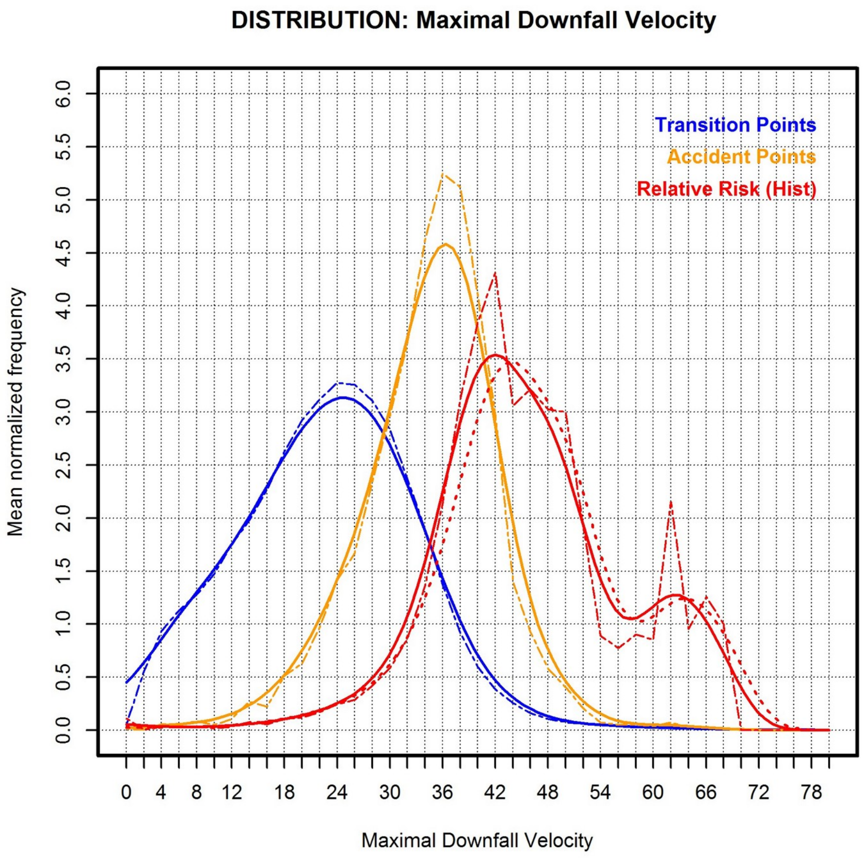
The relative risk continuously rises till a value of 60. Eventually values above 60 m/s2 are fatal to humans.

5.3 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]()		
Redundancy	FD_*		
Usage	***	Important property.	
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.



Interpretation

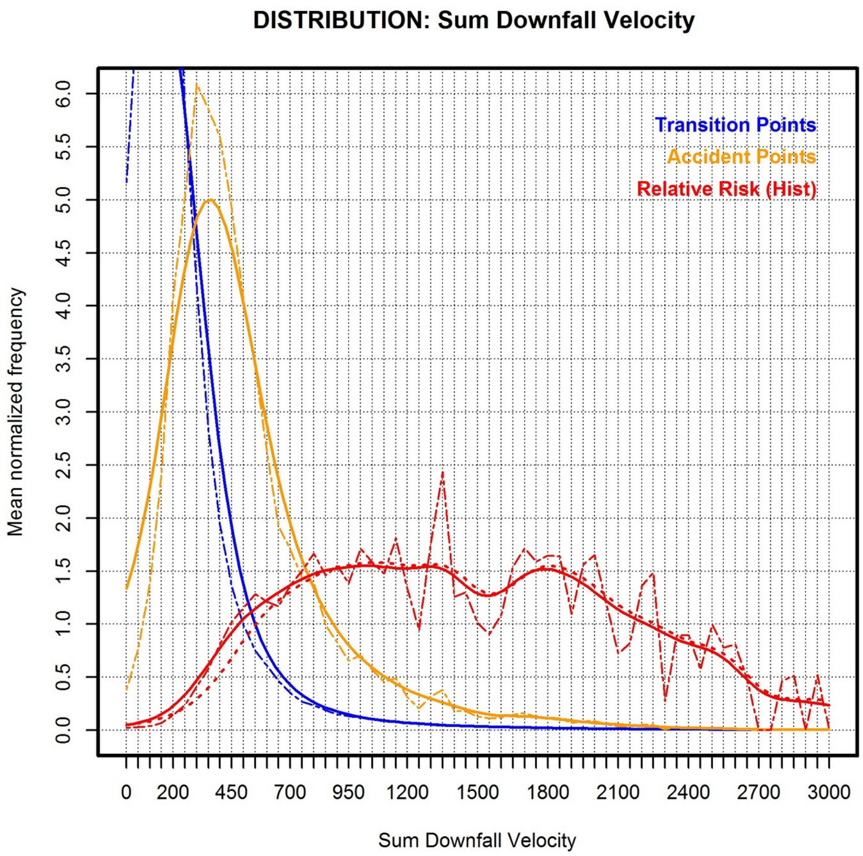
There is a constant rise of risk till the value of 42 m/s. Later on the data become uncertain.

5.4 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]		
Redundancy	FD_*		
Usage	**	Important property.	
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.



Interpretation

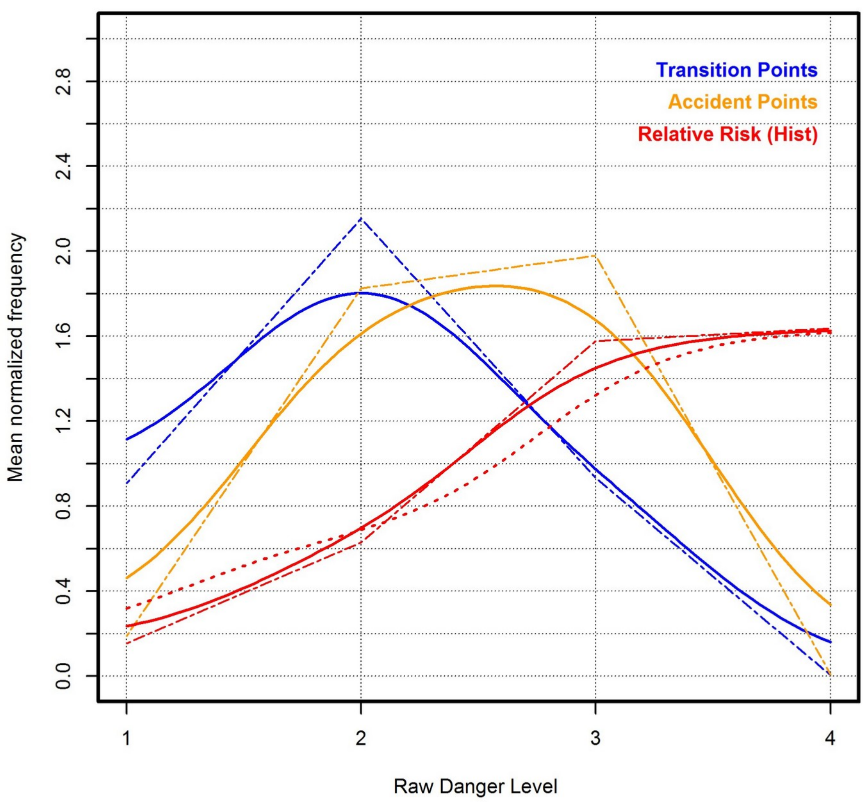
The relative risk continuously rises till a value of 800 and then remains stable.

6 Avalanche Forecast

6.1 Raw Danger Level (RDL)

Name	RDL	Raw Danger Level	
Description	Raw danger level according the avalanche forecast of the evening before.		
Comment			
Values	Integer	1, 2, 3, 4	-9999
Reference			
Redundancy	IDI, DI		
Usage	-	Use IDI, AOF, DCE resp. DI in stead.	
Copyrights	© SLF		

DISTRIBUTION: Raw Danger Level

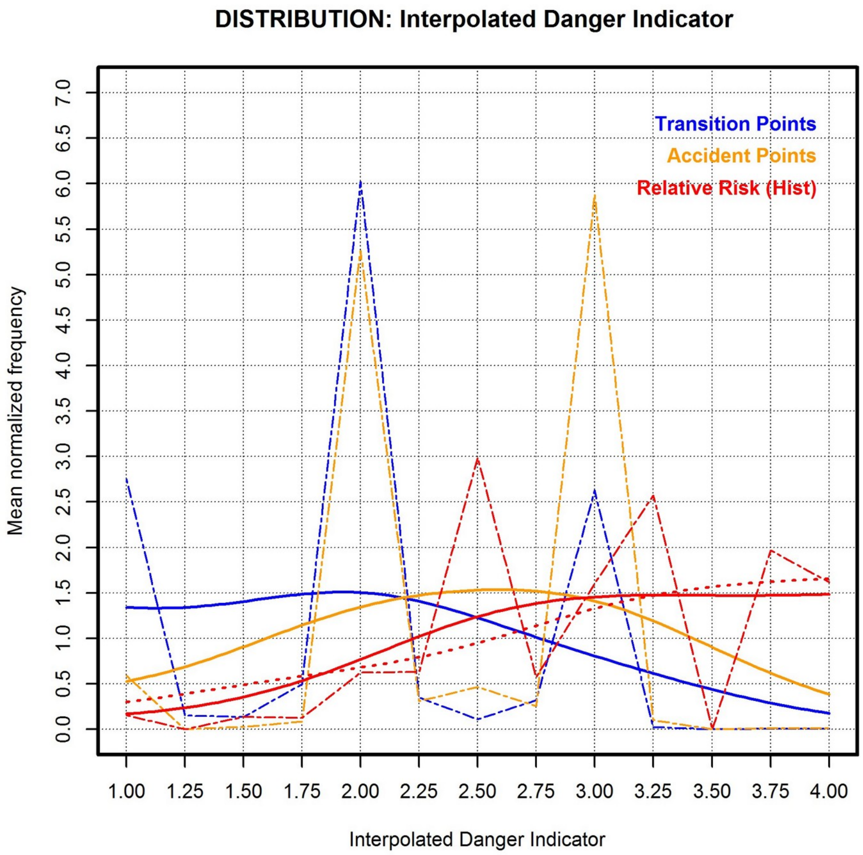


Interpretation

A stong rise in risk between 1 and 3. The risk at 4 is uncertain.

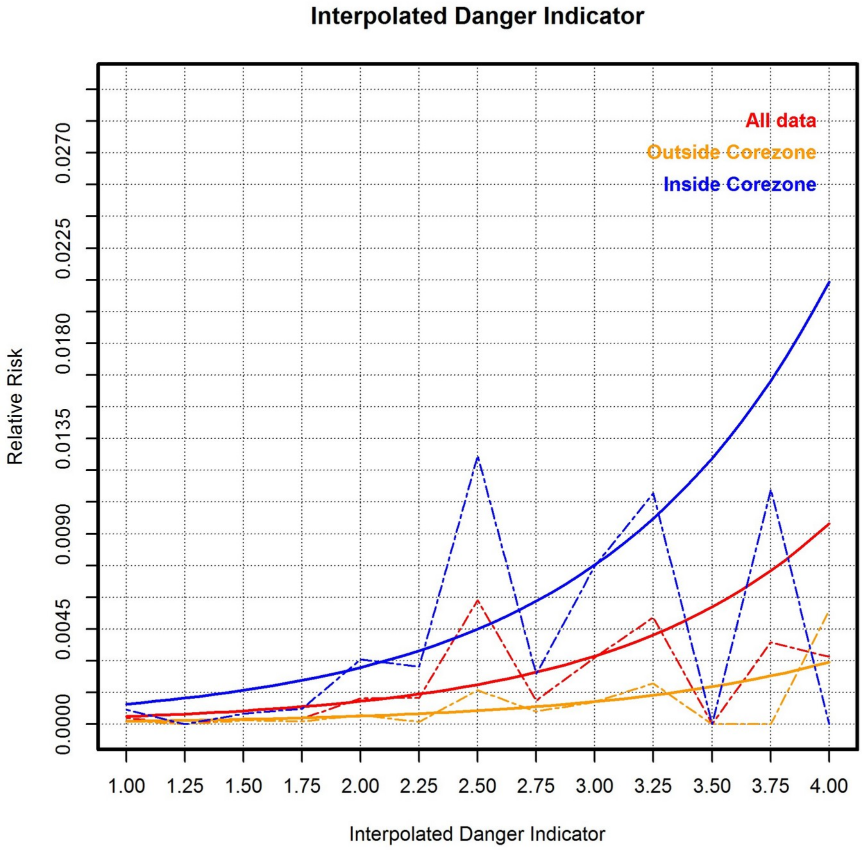
6.2 Interpolated Danger Indicator (IDI)

Name	IDI	Interpolated Danger Indicator	
Description	Horizontally interpolated danger indicator.		
Comment	Horizontal interpolation is calculated with up to 9 neighboring warning regions. In contrast to DI critical aspects and critical elevations are not taken into account. So the 1-level rule is **not** applied.		
Values	Decimal	1..4	-9999
Reference	Quantitative Risk Reduction Method (QRM), a data-driven avalanche risk estimator		
Redundancy	DI, RDL		
Usage	***	Important property, best used in combination with AOF and DCE.	
Copyrights	© Skitourenguru		



Interpretation:

There is a concentration of values at the integer values 1, 2, 3 and 4. Therefore we try to fit a statistical model. For binary discrete response variables typically a Binomial Generalized Linear Model (B-GLM) is fitted:



Interpretation

If we accept the assumption, the risk rises with the INV-LOGIT function, we can fit a binomial Generalized Model. The risk increases in the case of the red curve (all data) from danger level to danger level with the following factors: 2.96, 2.96, 2.95. On average, this makes a **factor of 2.96**.

The figure shows, that there is an important difference between the groups **inside the core zone** and **outside the core zone**. The horizontal shift between the blue curve (inside the core zone) and the yellow curve (outside the core zone) gives the x-value of the x-level rule. The value x is not constant, it rises slowly from left (1.94) to right (1.94). On average, it is 1.94. We should therefore rather apply a **1.94-level rule**.

Remark

Since V3.0.6 we only use backcountry skier accidents of the SLDB. Consequently we have very little accidents at danger level 4. Data become highly unreliable above danger level 3.

6.3 Critical Aspects (CA)

Name	**CA**	**Critical Aspects**	
Description	Critical aspects according the avalanche forecast of the evening before.		
Comment	Describes 8 sectors in clockwise direction, starting from the sector N-NNE. Example: 11100001 corresponds to NW over N to SE.		
Values	String	“XXXXXXXX”	“”
Reference			
Redundancy	AOF, CZ		
Usage	*	Use for subsampling.	
Copyrights	© SLF		

6.4 Critical Elevation (CE)

Name	**CE**	**Critical Elevation**	
Description	Critical elevation according the avalanche forecast of the evening before.		
Comment	If the value is positive, the dangerous elevations are above the indicated elevation. If the value is negative, the dangerous elevations are below the absolute value of the indicated elevation.		
Values	Decimal	-5000..+5000	-9999
Reference			
Redundancy	DCE, CZ		
Usage	*	Use for subsampling.	
Copyrights	© SLF		

6.5 Warning Region Code (WRC)

Name	**WRC**	**Warning Region Code**	
Description	Warning Region Code according the avalanche forecast of the evening before.		
Comment			
Values	Integer	0.9999	NA
Reference	Warning Regions according the SLF		
Redundancy			
Usage	*	Use for regional subsampling.	
Copyrights	© SLF		

6.6 Avalanche Problems (AP)

Name	AP	Avalanche Problems	
Description	Avalanche Problems according the avalanche forecast of the evening before.		
Comment	The bitmaps describes the avalanche problems with 6 bit in the following order: NEW_SNOW, WIND_SNOW, WET_SNOW, OLD_SNOW, GLIDE_SNOW, FAVOURABLE_SNOW, NO_DISTICT_PATTERN. Major and minor avalanche problems are not distinguished.		
Values	String	"XXXXXX"	""
Reference	Avalanche Problems according the SLF		
Redundancy			
Usage	*	Use for subsampling tests.	
Copyrights	© SLF		

The following table compares the risks between the avalanche problems:

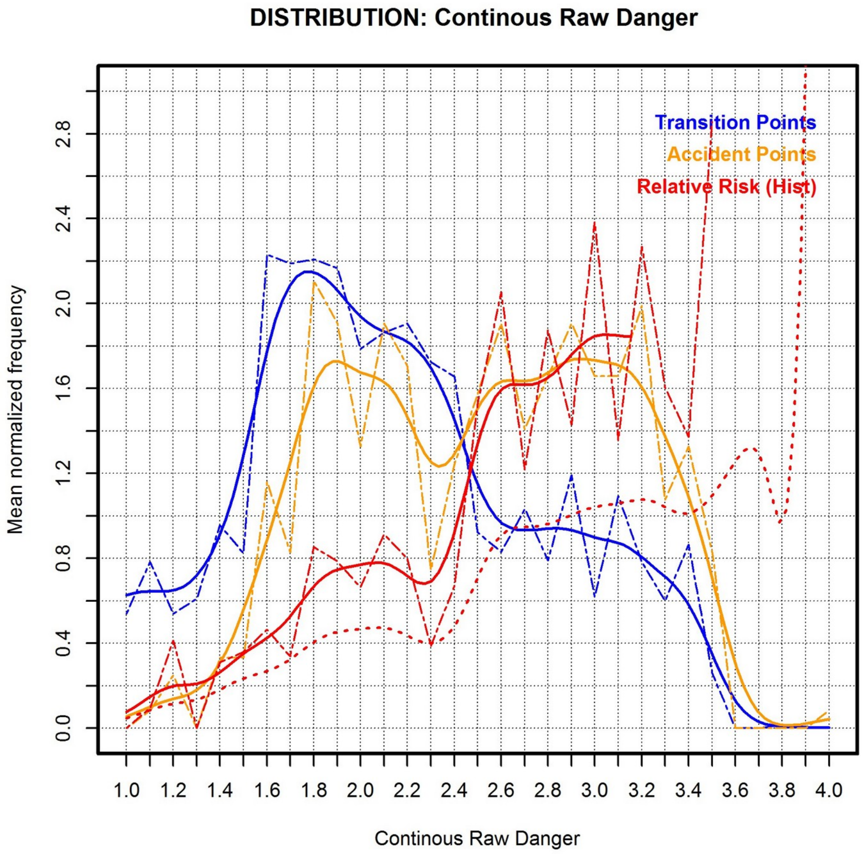
	NEW	WIND	WET	OLD	GLIDE
Accident Point Count (APC)	710	2737	924	1635	1176
Transition Point Count (TPC)	462208	2710812	976313	1870815	1433408
1000 * APC / TPC	1.54	1.01	0.95	0.87	0.82

Interpretation:

NEW_SNOW problem is twice as riskier then the other avalanche problems. Else there are no big differences between in risk between the avalanche problems. In particular the OLD_SNOW problem is not much riskier then the WE_SNOWT or WIND_SNOW snow problem.

6.7 Continous Raw Danger (CRD)

Name	CRD	Continous Raw Danger	
Description	A continuous danger value derived from RDL, CA, CE, AP and the texts (DE, EN, FR and IT) of the avalanche forecast		
Comment	Method based on text analytics and Gradient Boosting, that deduces a “continuous danger value” from the compiled text information (DE, EN, FR and IT) contained in the avalanche bulletin. Only avalanche forecasts from Winter 2013/14 to 2018/19 can be used, as standardized texts are not available in former versions of the avalanche forecast. Consequently 2/3 of the avalanche accidents can’t be covered.		
Values	Decimal	0..5	-9999
Reference	[To ski or not to ski, that is the question: Avalanche Risk Prediction with Text Analytics and Machine Learning]()		
Redundancy	DI, RDL		
Usage	**	Keep in mind the relative low number of available accidents during the time period.	
Copyrights	© Skitourenguru, © SAS, © SLF		



Interpretation

Travel usage and accidents have a clear inverse trend. Hence the risk rises significantly from CRD=1 to CRD=3. There are too few data for CRD>3.5. The data concentrations at the integer values can still be seen.

6.8 Distance to next lower danger level (DIST_LO)

Name	**DIST_LO**	**Distance to next lower danger level**	
Description	Distance from the point to the next warning region with a lower danger level.		
Comment	If 0, the distance couldn't be calculated. If negative, there wasn't found a lower danger level within the a search radius of abs(value).		
Values	Decimal	-∞..+∞	<=0
Reference			
Redundancy	IDI, DI		
Usage	*	Use in stead of IDI, DI	
Copyrights	© Skitourenguru		

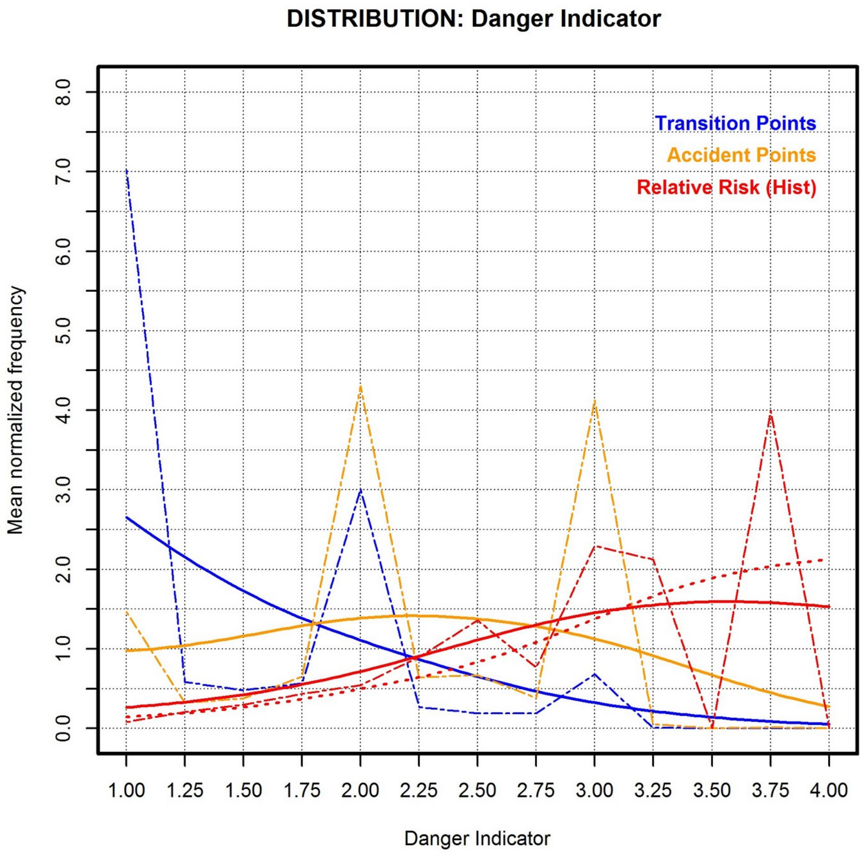
6.9 Distance to next higher danger level (DIST_HI)

Name	**DIST_HI**	**Distance to next higher danger level**	
Description	Distance from the point to the next warning region with a higher danger level.		
Comment	If 0, the distance couldn't be calculated. If negative, there wasn't found a higher danger level within the a search radius of abs(value).		
Values	Decimal	-∞..+∞	<=0
Reference			
Redundancy	ID, DI		
Usage	*	Use in stead of IDI, DI	
Copyrights	© Skitourenguru		

7 Avalanche Forecast and Terrain

7.1 Danger Indicator (DI)

Name	**DI**	**Danger Indicator**	
Description	The Danger Indicator takes into account RDL, Critical Aspects, Critical Elevations of the current warning region and of 9 neighboring warning regions. The avalanche forecast of the evening before is used.		
Comment	The 1-level rule is applied for the data of all involved warning regions.		
Values	Decimal	1..4	-9999
Reference	Quantitative Risk Reduction Method (QRM), a data-driven avalanche risk estimator		
Redundancy	RDL, IDI		
Usage	***	Important property, eventually use IDI, AOF and DCE in stead.	
Copyrights	© Skitourenguru		



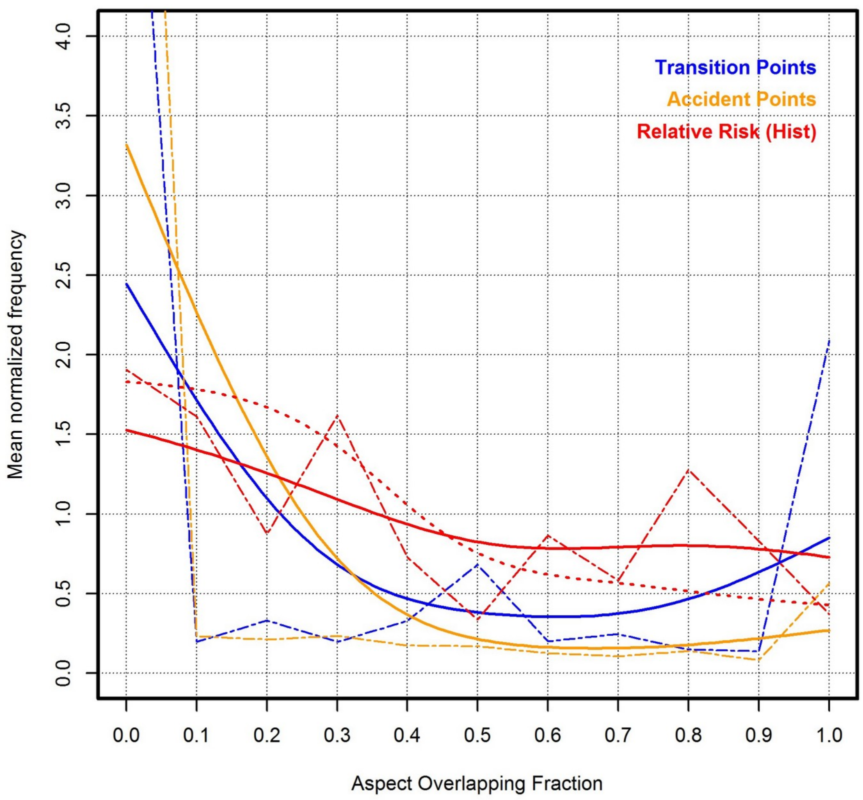
Interpretation

A rather linear rise in risk.

7.2 Aspect Overlapping Fraction (AOF)

Name	**AOF**	**Aspect Overlapping Factor**	
Description	The fraction of slope aspects, **that don’t overlap** with the critical aspects as indicated by the avalanche forecast of the evening before. Slope aspects are defined by the slope aspect range [MIN_ASPECT..MAX_ASPECT]. Only aspects with slope angle >25° are taken into account. If the slope aspect interval is undefined, the share of CA relative to 360° defines AOF.		
Comment	0 means the point (resp. the slope the point belongs to) is completely inside the critical aspects. 1 means the point (resp. the slope the point belongs to) is completely outside the critical aspects. If RDL=1 then CA=[-90°..135°] and CE=2000 is applied.		
Values	Decimal	0..1	-9999
Reference	Quantitative Risk Reduction Method (QRM), a data-driven avalanche risk estimator		
Redundancy	IDI, DI, RDL, CZ		
Usage	***	Important property.	
Copyrights	© Skitourenguru		

DISTRIBUTION: Aspect Overlapping Fraction



The next table shows the relative risk between three groups. All risks are expressed relative to a group that avoids totally the **critical aspects** (aof=1).

Dataset	aof = 0	(aof>0)&(aof<1)	aof ==1
All data	5.15	2.27	1.00
All data except rdl==1	4.32	2.23	1.00

The next table shows the relative risk between two groups. The risks are expressed relative to a group that avoids the **critical aspects** (aof>0.5).

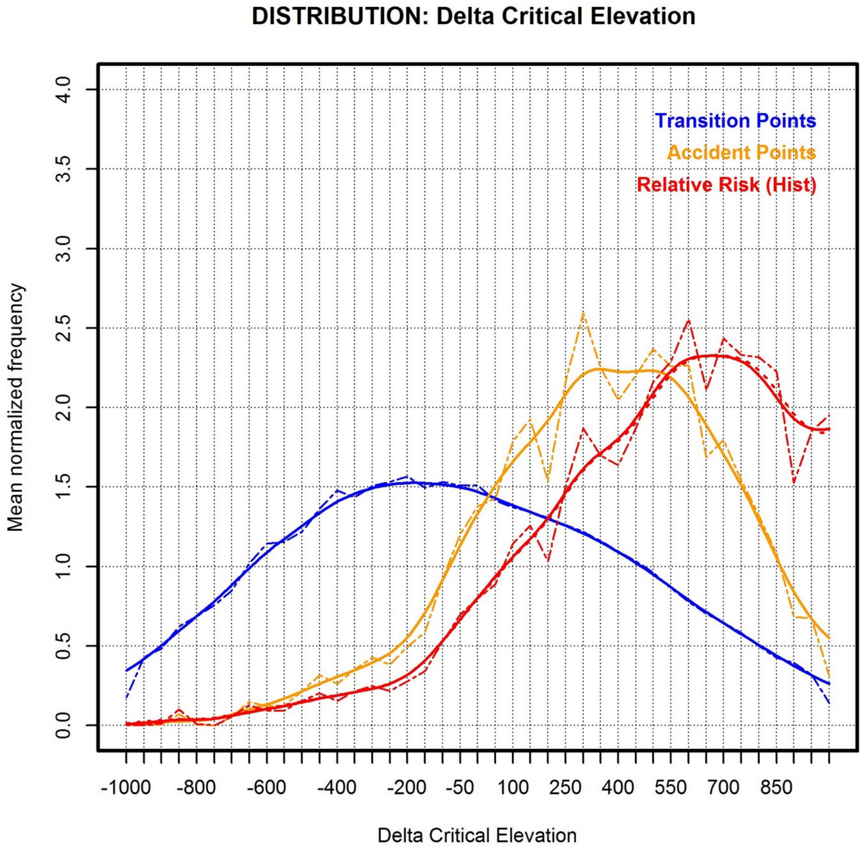
Dataset	aof < 0.5	aof > 0.5
All data	3.51	1.00
All data except rdl==1	3.09	1.00

Interpretation

The risk outside of the **critical aspects** is approximately 2-4 times lower then within the **critical aspects**.

7.3 Delta Critical Elevation (DCE)

Name	DCE	Delta Critical Elevation	
Description	Distinguish between two cases: • CE is positive: DCE = ELE - CE • CE is negative: DCE = abs(CE) - ELE Consequently a positive DCE always means the point is located inside the avalanche prone elevations. A negative DCE always means the point is located outside the avalanche prone elevations. If RDL=1 the critical elevation CE=2000 is always applied.		
Comment	Positive values mean the point is within the core zone. Negative values means the point is without the core zone.		
Values	Decimal	-1000..1000 m	-9999
Reference	Quantitative Risk Reduction Method (QRM), a data-driven avalanche risk estimator		
Redundancy	IDI, DI, RDL, CZ		
Usage	***	Important property.	
Copyrights	© Skitourenguru		



The next table shows the relative risk between two groups. The risks are expressed relative to a group that avoids the **critical elevations** (dce<0).

Dataset	dce>0	dce<0
All data	6.93	1.00
All data except rdl==1	7.80	1.00

Interpretation

There is a very significant rise in risk in function of DCE. Data below -600 and above +600 m are uncertain. Interestingly the factor **critical elevation** seems to be more significant then the factor **critical aspects**.

7.4 Core Zone (CZ)

Name	CZ	Core Zone	
Description	Is the point within the core zone (1) or outside the core zone (0)		
Comment	CZ = ((AOF < 0.5) && (DCE > 0))? 1:0 If RDL=1 the critical aspects CA=[-90°..90] are always applied. If RDL=1 the critical elevation CE=2000 is always applied.		
Values	Integer	0, 1	NA
Reference	Quantitative Risk Reduction Method (QRM), a data-driven avalanche risk estimator		
Redundancy	IDI, DI, RDL, CZ		
Usage	***	Use for subsampling.	
Copyrights	© Skitourenguru		

The property allows to calculate the effect of the core zone on the risk of two winter sportsmen: Sportsman A is always en route outside of the core zone, sportsman B is always en route inside the core zone. By simple row counting we can conclude **sportsman B has 7.43 time higher risk then sportsman A**. If we omit all data with rdl=1, **sportsman B has a 7.44 time higher risk then sportsman A**

Now we can further refine the result and check the effect of AOF and DCE. The following table will calculate relative risk indicators. A sportsman who is outside of the core zone in a double sense (relative to the **critical aspects** and the **critical elevations**) has the relative risk 1.

	Inside (AOF< 0.5)	Outside (AOF> 0.5)
Inside (DCE> 0)	32.88	8.93
Outside (DCE < 0)	4.64	1.00

Interpretation

The effect of the core zone is quite important, where the effect of **elevation** is more or less twice as important then the effect of **aspect**.

8 Spatio-temporal information

8.1 Date (DATE)

Name	DATE	Date	
Description	The date the accident occurred (failure data) resp. the date the point was passed by a backcountry skier (success data).		
Comment	The information of the avalanche bulletin comes from the day before (evening forecast 17 h)		
Values	String	"yyyy.MM,dd"	NA
Reference			
Redundancy	HASH		
Usage	*	Use for temporal subsampling.	
Copyrights	© Skitourenguru		

8.2 X- Coordinate (X)

Name	X	X-Coordinate	
Description	The X-Coordinate of the accident point (failure data) resp. the transition point (success data).		
Comment	The coordinate is expressed in EPSG=21781.		
Values	Decimal	400'000..900'000	NA
Reference			
Redundancy	HASH		
Usage	*	Use for regional subsampling.	
Copyrights	© Skitourenguru		

8.3 Y- Coordinate (Y)

Name	Y	Y-Coordinate	
Description	The Y-Coordinate of the accident point (failure data) resp. the transition point (success data).		
Comment	The coordinate is expressed in EPSG=21781.		
Values	Decimal	50'000..300'000	NA
Reference			
Redundancy	HASH		
Usage	*	Use for regional subsampling.	
Copyrights	© Skitourenguru		

8.4 Hash (HASH)

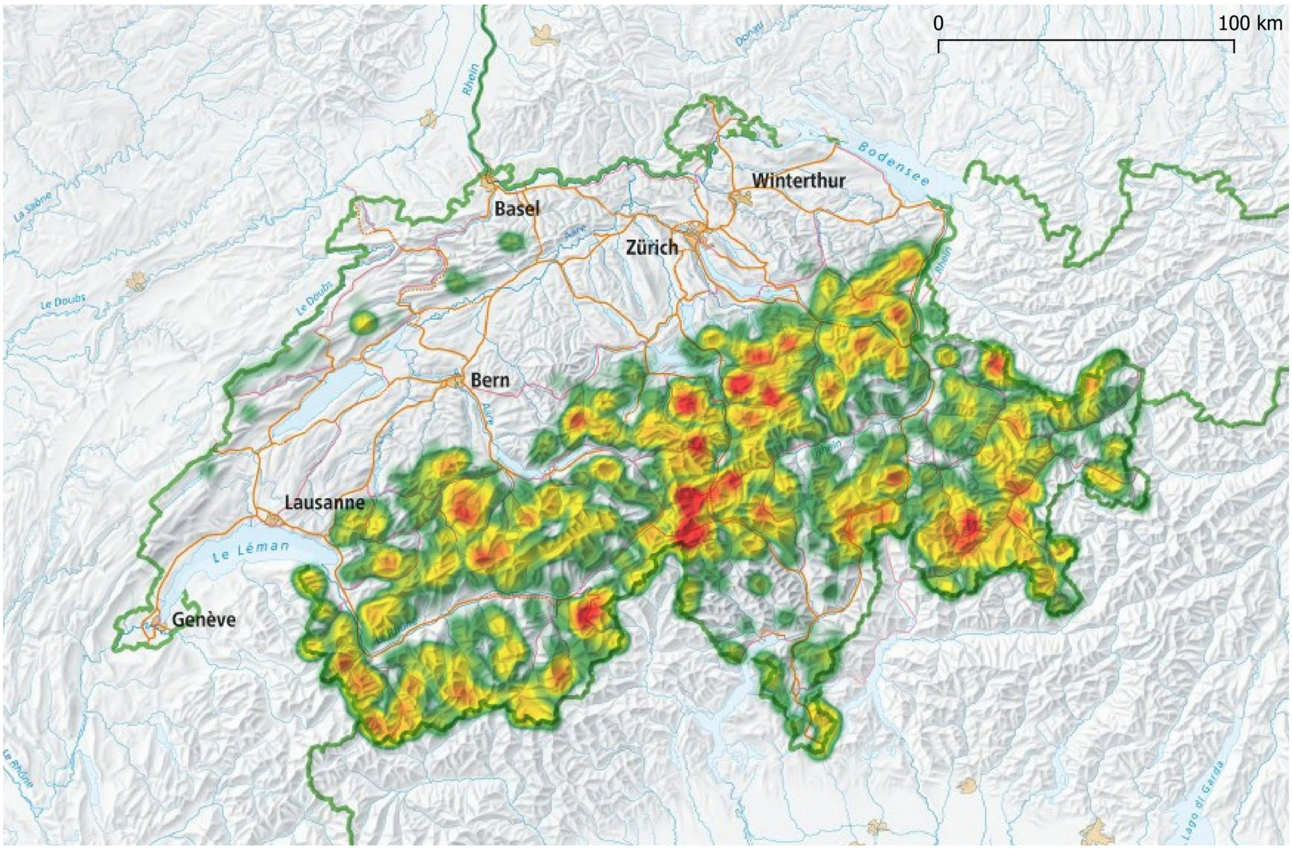
Name	**HASH**	**Hash**	
Description	A SHA256 hash over DATE, X and Y.		
Comment	Identifies uniquely a point in space and time.		
Values	String	SHA256 hash	NA
Reference			
Redundancy	DATE, X, Y		
Usage	-	-	
Copyrights	© Skitourenguru		

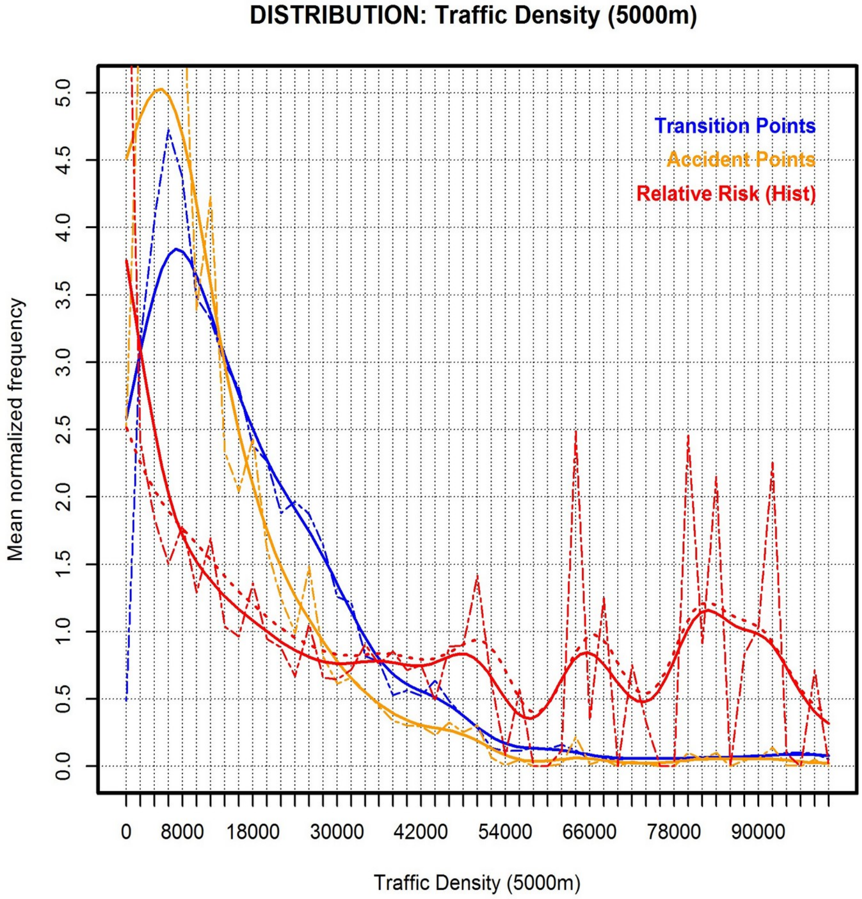
9 Human-related Information

9.1 Traffic Density (TD5000)

Name	TD5000	Traffic Density	
Description	Describes the overall backcountry skier traffic density at the point (heatmap).		
Comment	The heatmap was calculated from the GPS tracks collection. The kernel bandwidth is 5000 m, which leads to a generalized overall travel density.		
Values	Decimal	0..125’000	-9999
Reference	Heatmap of Skitourenguru		
Redundancy	DIST_SAC, DIST_PISTE, TD100		
Usage	*	Use with low priority, as data is only available for Switzerland. Use for subsampling.	
Copyrights	© Skitourenguru		

The following figure shows the current valid heatmap (V3.0) based on GPS tracks collected till June 2019:





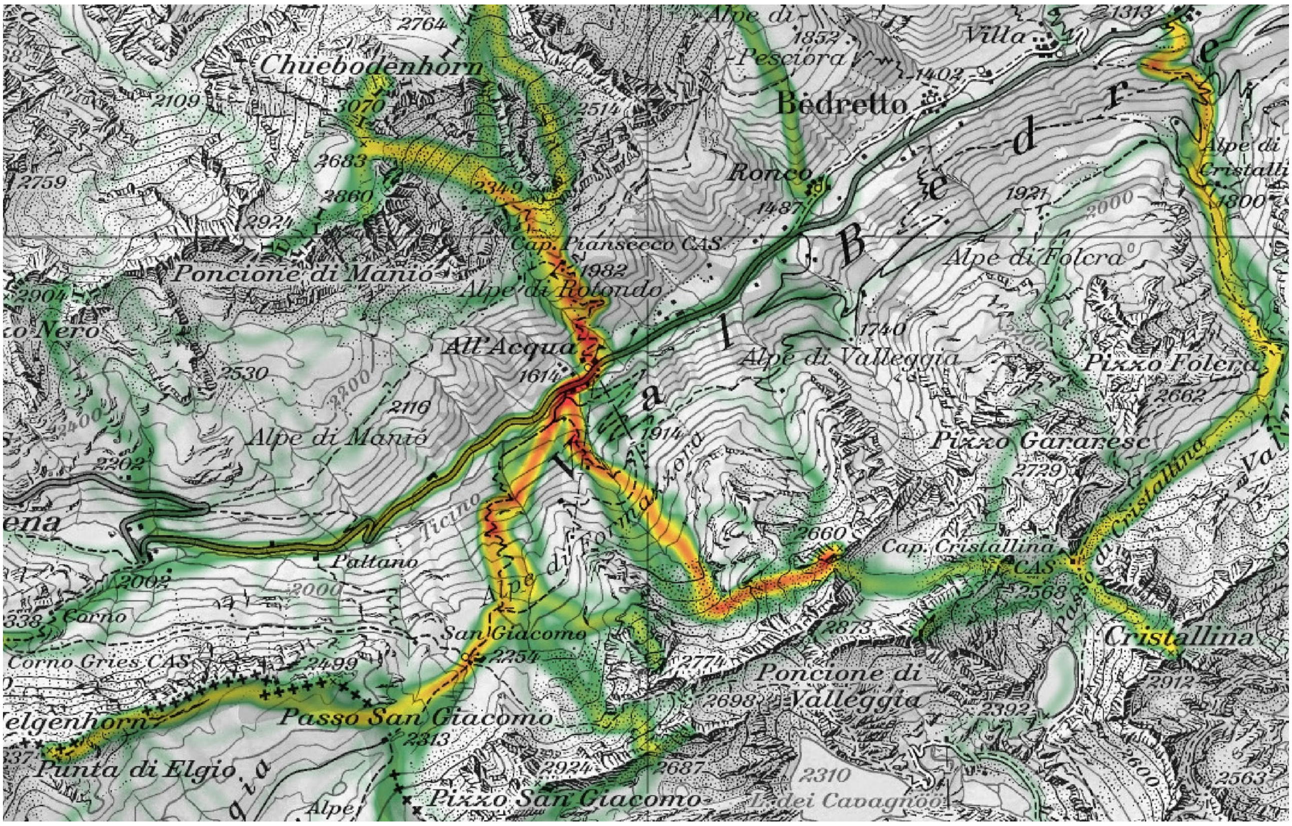
Interpretation

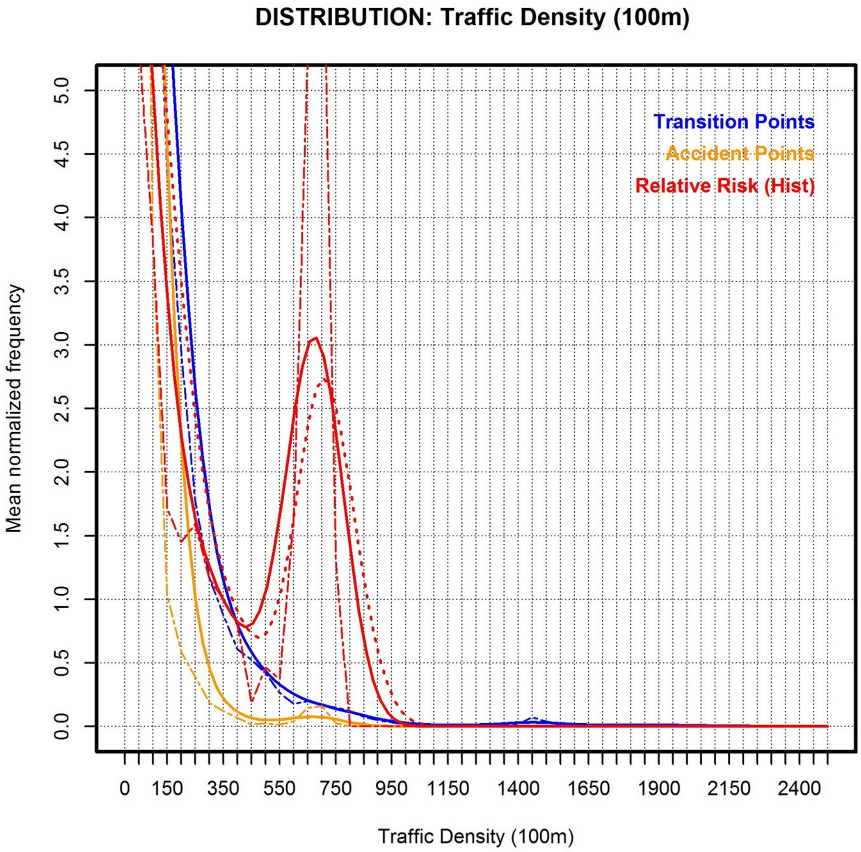
The relative risk constantly decreases with increasing overall traffic. Data below 1500 are uncertain. Be careful with interpretation: TD5000 is calculated from the same data set as already used for the success points.

9.2 Traffic Density (TD100)

Name	**TD100**	**Traffic Density**	
Description	Describes the local backcountry skier traffic density at the point (heatmap).		
Comment	The heatmap was calculated from the GPS tracks collection. The kernel bandwidth is 100 m, which leads to a local travel density (corridor).		
Values	Decimal	0..125’000	-9999
Reference	[Heatmap of Skitourenguru]		
Redundancy	DIST_SAC, DIST_PISTE, TD5000		
Usage	*	Use with low priority, as data is only available for Switzerland. Use for subsampling.	
Copyrights	© Skitourenguru		

The following figure shows the current valid heatmap (V3.0) based on GPS tracks collected till June 2019:



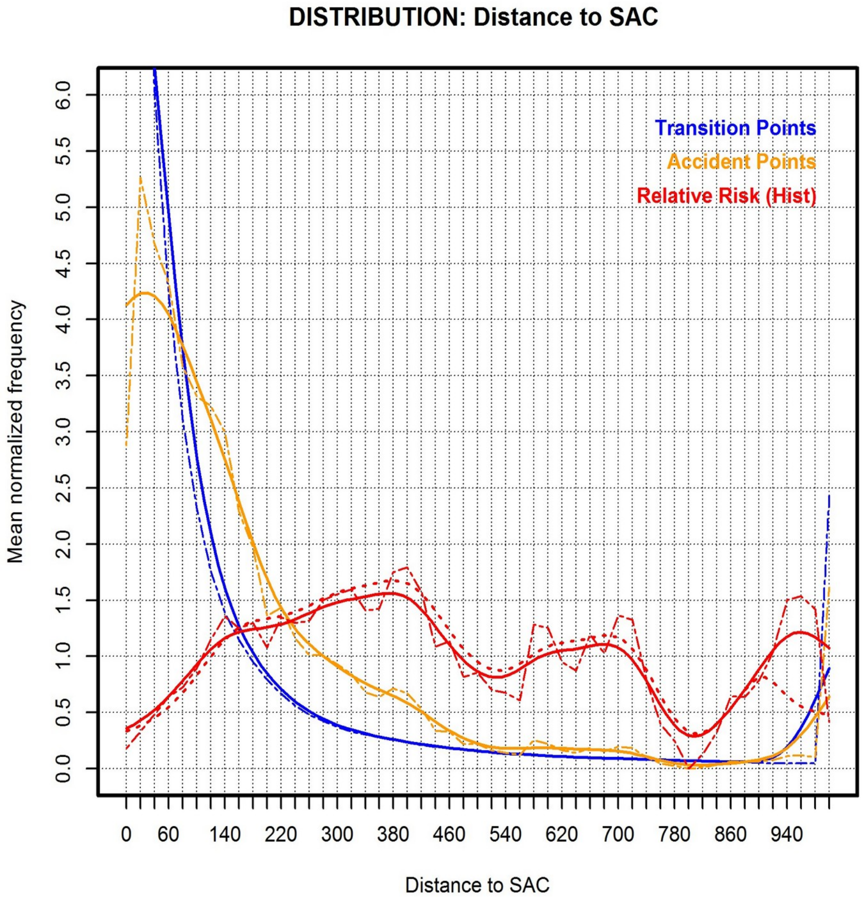


Interpretation

Be careful with interpretation: TD100 is calculated from the same data set as already used for the success points.

9.3 Distance to next SAC skitour (DIST_SAC)

Name	DIST_SAC	Distance to next SAC skitour	
Description	The distance to the next skitour of the Swiss Alpine Club (SAC) backcountry skiing network.		
Comment	All point with a distance larger then 1000 m will have the value 1000 m.		
Values	Decimal	0..1000 m	1000 m
Reference	Backcountry Skiing Maps of Switzerland (Swisstopo/SAC)		
Redundancy	TD, DIST_PISTE		
Usage	*	Use with low priority, as data is only available for Switzerland. Use for subsampling.	
Copyrights	© Swisstopo, SAC		

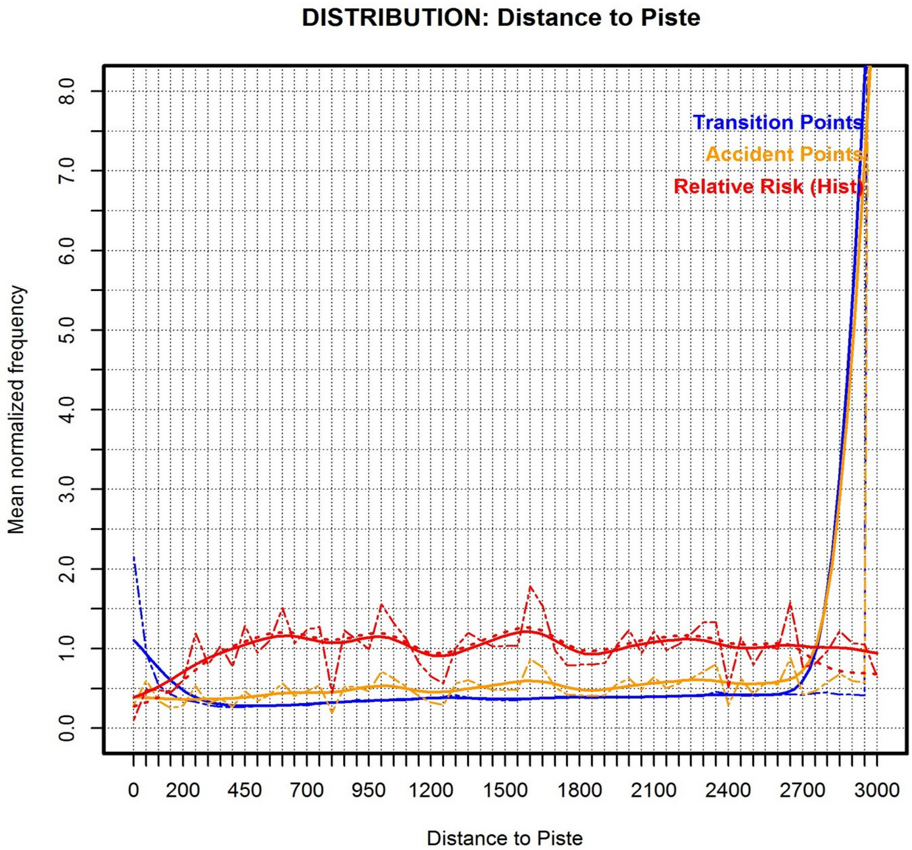


Interpretation

There is a constant rise in risk with rising distance to the next SAC skitour. Above 500 m the result is unreliable.

9.4 Distance to next piste (DIST_PISTE)

Name	DIST_PISTE	Distance to next SAC skitour	
Description	The distance to the next downhill piste		
Comment	All point with a distance larger then 3000 m will have the value 3000 m		
Values	Decimal	0..3000 m	3000 m
Reference	OSM tag Piste		
Redundancy	TD, DIST_SAC		
Usage	-	Use only for subsampling: With this property its possible to filter out accidents that occurred in a freeriding context.	
Copyrights	© OSM		



Interpretation

There is no clear trend.

9.5 Identifier of Route (ID)

Name	**ID**	**Identifier of Route**	
Description	The identifier of the route.		
Comment	In case of accidents its the ID of the accident. In case of GPS tracks its an ID of the route.		
Values	String		
Reference			
Redundancy			
Usage	-	Use for subsampling or statistical tests	
Copyrights	© Skitourenguru		

9.6 Elevation Gain (EG)

Name	**EG**	**Elevation Gain of the Route**	
Description	The elevation gain of the route the point is member of.		
Comment	In case of accidents the elevation gain will be always around 40 m and has no particular meaning. In case of the GPS tracks its the elevation gain of the route the point is member of.		
Values	Decimal	0..10’000 m	NA
Reference			
Redundancy			
Usage	-	Use for subsampling	
Copyrights	© Skitourenguru		

10 Recommendations

10.1 Multivariate Regression Analysis

Use the following properties as explanatory variables:

1. Terrain (Avalanche probability): TI (alternatively MSA*, HP or SP), ELE, FOLD, FD, DIST_RIDGE
2. Terrain (Avalanche consequences): FD_MAXV, FD_SUMV
3. Avalanche Forecast: DI (or alternatively IDI, AOF and DCE)
4. Human related Information: -

10.2 R-Statistics

Be careful when loading the GPS dataset. R-Statistics has problems to load all rows with all columns. Load the columns selectively with read.csv (colClasses=..). Selective loading has the advantage to be fast.