Terrain Usage Dataset (TUD)

User Manual

Günter Schmudlach

11. July 2022

(V1.0.3)

Abstract

The document introduces a training dataset, that enables the understanding of the t**errain usage** of backcountry skiers. The dataset contains raster data about **travel frequency** and raster data about **terrain properties**. The dataset can be used to train a model able to infer **travel frequency** from **terrain properties**.

Content

1 Introduction

Backcountry skiers are selective when choosing terrain to ascend and descend with skies. For example its known that in the Alps backcountry skiers have a preference for terrain with a slope angle around 20°.

Since a few years algorithms that draw automatically **backcountry skiing maps** are under development.[1] These algorithms calculate minimal cost paths trough a **cost surface**. The better the cost surface represents the preferences of backcountry skiers, the better the resulting **backcountry skiing maps**. In order to develop good cost surfaces its crucial to model the terrain preference of backcountry skiers.

With the omnipresence of GPS devices among the skier community a lot of data is collected about the terrain usage of backcountry skiers. From these data it is possible to derive **travel frequency** maps in a high-resolution raster. Such maps are also known under the name “heatmap”.

High-resolution data about the terrain make it possible to derive a series of **terrain properties** like slope angle, forest density and many more**.** With methods from “predictive modeling” its possible to understand whether there is a link from the terrain properties to the travel frequency:

travelFreqeuncy = f(terrainProperties)

Predictive modeling has the potential to infer the function f from a training dataset.

2 Data description

2.1 Format

The dataset contains OGC-compliant raster data in the XYZ format or in the GTiff format. The raster data can be visualized with Geographical Information Systems (GIS) like qGis or ArcGis. The metadata about all raster dataset are:

- Extent: The data covers three different extents in Switzerland (tiny, small and large).
- Projection: Coordinates are expressed with the projection EPSG=21781.
- Resolution: 10 m x 10 m

Its also possible to load the data into R-Statistics, if GDAL is installed: Raster Data in R.

As the dataset is relatively big calculation times can become slow. Therefore the dataset contains 3 extents:

1. **Tiny**: An extent of 4 km x 4 km in the area of Gotthard. If the data is delivered in the XYZ format the csv files can be opened with excel or a text editor. Use this extent only for quick data inspection. Don’t use the extent to train models, it doesn’t contain enough knowledge.
2. **Small**: An extent of 25 km x 25 km in the area of Gotthard. Use this extent for fast model development.
3. **Large**: An extent around Switzerland. Use this extent for the final model. In order to analyze only data from the Alps, apply a spatial filter with the property **FILTER** (see **4.14**).

2.2 Data filter modes

The described data contains two sets about the travel frequency:

1. A heatmap calculated from a huge collection of GPS tracks processed by Skitourtenguru (see property **TD_X**). TD stands for Travel Density, X is a Kernel Radius used for smoothing.
2. The heatmap of Strava (see property **STRAVA**)

Both travel frequency maps show no travel activity in many spots. On such spots the travel frequency maps hold the no-data value 0. Obviously that doesn’t mean, that such spots don’t have terrain properties for good skiing. Therefor it makes sense to develop models with two data filter modes:

1. All Data Model (ADM): All data (including 0 values) are used as training data.

1 A. Eisenhut: Automatisch generierte Skitouren

2. Unequal to Zero Data Model (UZDM): Only data with values >0 are used as training data.

The model scores will show which models have the better performance.

2.3 Data matching

In order to process the data some caution regarding **data matching** and handling of **no-data** values must be taken:

1. Read all single properties.
2. Match the properties into one table. Use x and y to match the properties. x and y in conjunction are unique. Make sure all rows could be matched.
3. **fd_risk** and **ti**: Replace the value -9999 with 0.
4. Make sure the table doesn't contain anymore the value -9999.
5. Make sure that no table contains non-number values (like NaN).
6. If **strave** is used, only take values where ele>1500. Below 1500 m the raster can contain many errors.

2.4 Data description scheme

Each raster property of the dataset is described with a simple table as follows:

Name	**Abbreviation**	**Name**	
Description	Description of the property.		
Comment	A comment about the property.		
Values	Data type	Value range	No data value ([2])
Reference	Reference to more information about the property.		
Usage	0-3 Stars	Recommendations for the usage of the property.	
Copyrights	Copyrights of the raw data		

2 YES, if the raster effectively holds no-data values, else NO.

3 Dependent variables

3.1 Travel density (TD_10)

Name	TD_10	Travel density 10m	
Description	The travel density derived from a collection of GPS tracks and smoothed with a Kernel of 10 m. Calculated with v.kernel and multiplier=35900. Thanks to the multiplier the mean is approximately 100.		
Comment	The value is on a “Ratio scale”		
Values	Decimal	0..62852	0 (YES)
Reference	On the correlation between the forecast avalanche danger and avalanche risk taken by backcountry skiers in Switzerland		
Usage	*	Use TD_50	
Copyrights	© Skitourenguru		

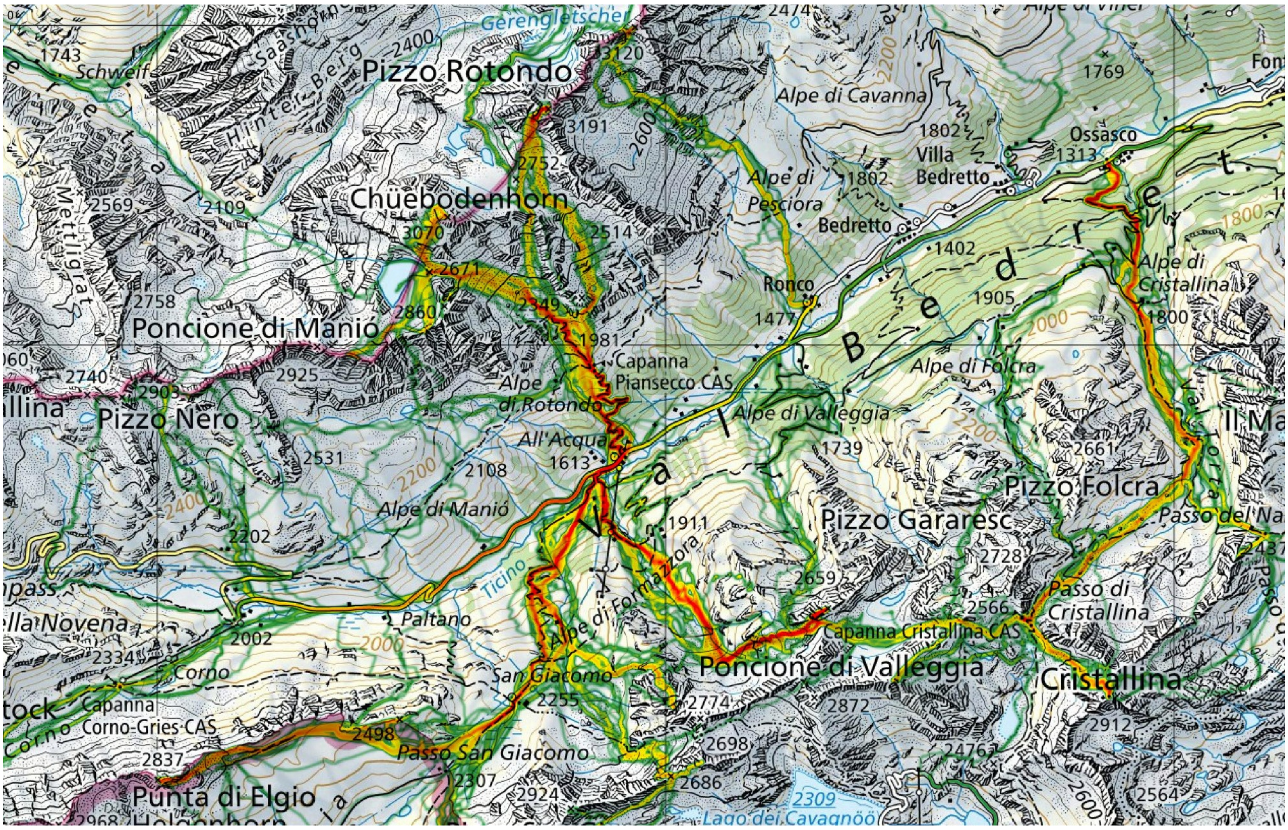
3.2 Travel density (TD_20)

Name	TD_20	Travel density 20m	
Description	The travel density derived from a collection of GPS tracks and smoothed with a Kernel of 20 m. Calculated with v.kernel and multiplier=35900. Thanks to the multiplier the mean is approximately 100.		
Comment	The value is on a “Ratio scale”		
Values	Decimal	0..?	0 (YES)
Reference	On the correlation between the forecast avalanche danger and avalanche risk taken by backcountry skiers in Switzerland		
Usage	*	Use TD_50	
Copyrights	© Skitourenguru		

3.3 Travel density (TD_50)

Name	TD_50	Travel density 50m	
Description	The travel density derived from a collection of GPS tracks and smoothed with a Kernel of 50 m. Calculated with v.kernel and multiplier=35900. Thanks to the multiplier the mean is approximately 100.		
Comment	The value is on a “Ratio scale”		
Values	Decimal	0..?	0 (YES)
Reference	On the correlation between the forecast avalanche danger and avalanche risk taken by backcountry skiers in Switzerland		
Usage	***	Use TD_50	
Copyrights	© Skitourenguru		

The following image shows an example. The traffic light colors indicate the travel frequency on the particular spot (10x10 m - Pixel):



3.4 Travel density (TD_100)

Name	**TD_100**	**Travel density 100m**	
Description	The travel density derived from a collection of GPS tracks and smoothed with a Kernel of 100 m. Calculated with v.kernel and multiplier=35900. Thanks to the multiplier the mean is approximately 100.		
Comment	The value is on a “Ratio scale”		
Values	Decimal	0..?	0 (YES)
Reference	On the correlation between the forecast avalanche danger and avalanche risk taken by backcountry skiers in Switzerland		
Usage	*	Use TD_50	
Copyrights	© Skitourenguru		

3.5 Strava (STAVA)

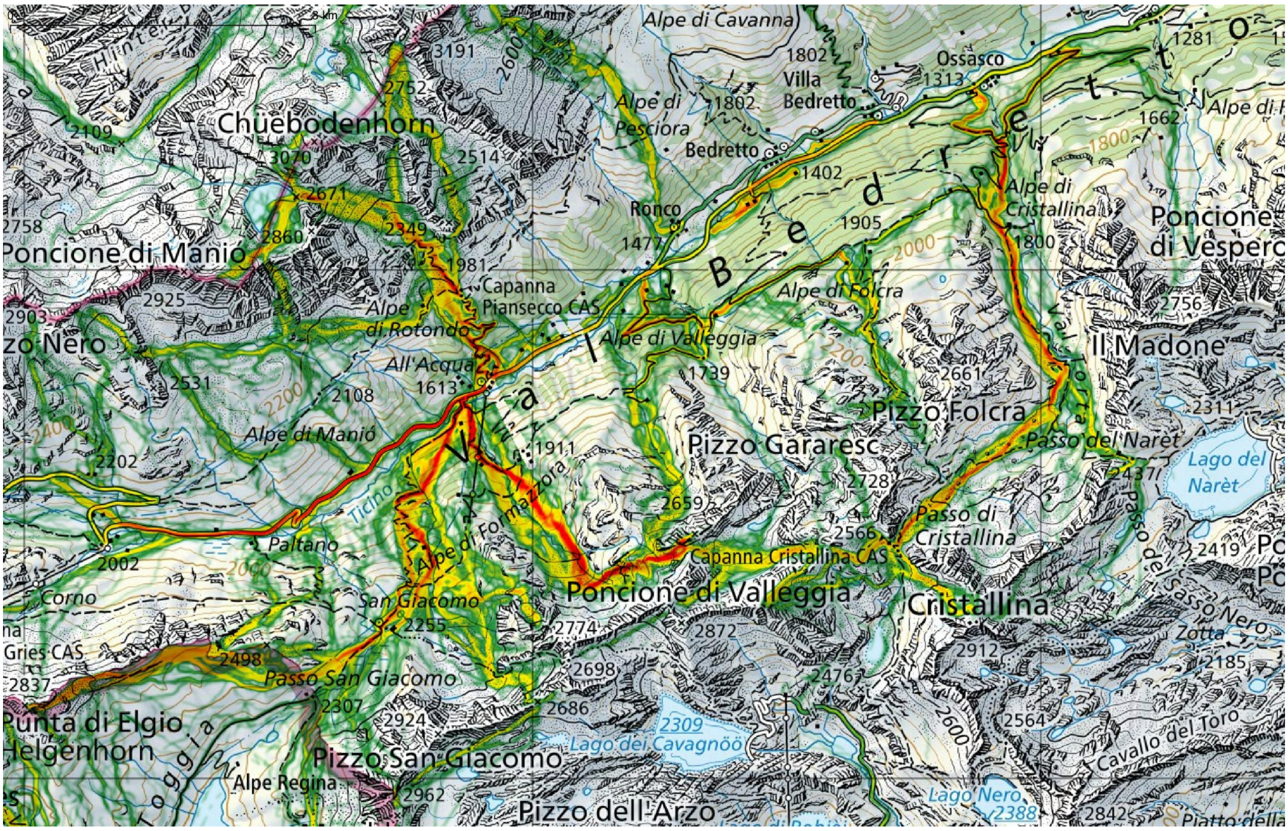
Name	STAVA	Heatmap from Strava	
Description	A heatmap from Strava covering approximately the years 2016..2018 and 2020..2022.		
Comment	The value fulfills only trichtonomy and transitivity.		
Values	Decimal	0..1518	0 (YES)
Reference	Strava		
Usage	***	The property is only interesting for internal comparison.	
Copyrights	© Strava (Due to ownership of the data no results that include this property can be published)		

Little is known about the pre-processing procedure and the updating strategy of Strava.

Strava heatmaps are published with RGB values. The present raster is the result of the additionof the trhree RGB bands. Subsequently the raster is smoothed with a Gauss-Kernel of width 25 m. Ski piste and cablecars are nulled. Not everywhere the results are perfect. Particularly below 1500 m there can be a lot of wrong data (for instance nordic walking piste can't be nulled everywhere). **So eventually process only spots, where ele>1500m**.

In comparison to TD_X this property is based on much more GPS tracks.

The following image shows an example. The traffic light colors indicate the travel frequency on the particular spot (10x10 m – Pixel):



4 Independent variables

4.1 Slope Angle (SLOPE)

Name	**SA**	**Slope Angle**	
Description	The slope angle derived from a DEM with 10 m resolution.		
Comment			
Values	Decimal	0..90°	-9999 (NO)
Reference	gdaldem (slope)		
Usage	***		
Copyrights	© Swisstopo		

4.2 Plan Curvature (PLANC7)

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). The property indicates if a spot is located on a ridge, in a valley or on a homogeneous slope.		
Values	Decimal	-100..100	-9999 (NO)
Reference	r.param.scale(size=7, method=planc)		
Usage	**		
Copyrights	© Skitourenguru		

4.3 Forest Density (FD)

Name	**FD**	**Forest Density**	
Description	Forest Density (in %) and a resolution of 10 m.		
Comment			
Values	Decimal	0..100%	-9999 (NO)
Reference	Tree Cover Density (2018)		
Usage	**		
Copyrights	© ESA		

4.4 Fall down risk (FD_RISK)

Name	FD_RISK	Fall down risk	
Description	The product of SLOPE (Slope angle) and FD_MAXV (maximal velocity on down fall trajectory).		
Comment	SLOPE can be interpreted as a proxy for the probability to fall down. FD_MAXV can be interpreted as the consequences of falling 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 (YES)
Reference			
Usage	***	The no-data value -9999 means that there is no risk to fall down. Therefore -9999 should be replaced by 0.	
Copyrights	© Skitourenguru		

4.5 Elevation (ELE)

Name	ELE	Elevation	
Description	Elevation according to the DEM with 10 m resolution.		
Comment			
Values	Decimal	0..5000 m	-9999 (NO)
Reference	swissALTI3D-10m		
Usage	***		
Copyrights	© Swisstopo		

4.6 Terrain Indicator (TI)

Name	TI	Terrain Indicator	
Description	TI indicates how suitable a terrain point is to trigger an avalanche (MRSAR=100 m).		
Comment			
Values	Decimal	0..1	-9999 (YES)
Reference	Method for an Automatized Avalanche Terrain Classification		
Usage	**	The no-data value -9999 means that there is no avalanche risk Therefore -9999 should be replaced by 0.	
Copyrights	© Skitourenguru		

4.7 Distance to next Swiss Alpine Club Skitour (SAC_DIST)

Name	SAC_DIST	Distance to next Swiss Alpine Club Skitour	
Description	Distance in “Meters” to the next skitour according to the ski tours of the Swiss Alpine Club (SAC).		
Comment			
Values	Decimal	0..16101	-
Reference	Swiss Alpine Club Ski tours on map.geo.admin.ch		
Usage	**	The property can be used as spatial filter for the dataset **large** and substituting the property **FILTER**. A threshold of 3000 m is a reasonable cut-off for filtering ski touring terrain.	
Copyrights	© Swiss Alpine Club, Swisstopo, Skitourenguru		

4.8 Identifier of nearest Swiss Alpine Club Skitour (SAC_ID)

Name	SAC_ID	Identifier of nearest Swiss Alpine Skitour	
Description	The identifier (id) to the nearest skitour of the Swiss Alpine Club (SAC)		
Comment			
Values	Integer	0..17836	-
Reference	Swiss Alpine Club Ski tours on map.geo.admin.ch		
Usage	**	The id can be used to represent the effect of route popularity	
Copyrights	© Swiss Alpine Club, Swisstopo, Skitourenguru		

4.9 Distance to next Skitourenguru Skitour (STG_DIST)

Name	STG_DIST	Distance to next Skiturenguru Skitour	
Description	Distance in “Meters” to the next skitour of Skitourenguru.		
Comment			
Values	Decimal	0..?	-
Reference	https://www.skitourenguru.com		
Usage	**		
Copyrights	© Skitourenguru		

4.10 Identifier of nearest Skitourenguru Skitour (STG_ID)

Name	**STG_ID**	**Identifier of nearest Skitourenguru Skitour**	
Description	The identifier (id) to the nearest skitour of Skitourenguru		
Comment			
Values	Integer	0..3299	-
Reference	https://www.skitourenguru.com		
Usage	**	The id can be used to represent the effect of route popularity	
Copyrights	© Skitourenguru		

4.11 Highway (HIGHWAY)

Name	**HIGHWAY**	**Highway**	
Description	Category, that indicates distance to the next street: • 1: Next street in distance 0...5 m • 2: Next street in distance 5...15 m • 3: Next street in distance 15...25 m • 4: Next street in distance 25...35 m • 0: No street nearby		
Comment	A street is defined as a roadway that can be managed by an agricultural vehicle. The name “highway” comes from the OSM tag “highway”.		
Values	Binary	0,1,2,3,4	
Reference			
Usage	**	Consider 1,2,3 on the street.	
Copyrights	© OSM		

4.12 Path (PATH)

Name	**PATH**	**Path**	
Description	Category, that indicates distance to the next footpath: • 1: Next footpath in distance 0...5 m • 2: Next footpath in distance 5...15 m • 3: Next footpath in distance 15...25 m • 4: Next footpath in distance 25...35 m • 0: No footpath nearby		
Comment	A footpath usually can’t be managed anymore by an agricultural vehicle.		
Values	Binary	0,1,2,3,4	none
Reference			
Usage	**	Consider 1,2 on the footpath.	
Copyrights	© OSM		

4.13 Lake (LAKE)

Name	**LAKE**	**Lake**	
Description	Indicates if the pixel is located on a lake (1) or on land (0).		
Comment			
Values	Binary	0,1	none
Reference			
Usage	**		
Copyrights	© OSM		

4.14 Filter (FILTER)

Name	**FILTER**	**Lake**	
Description	Indicates if the pixel is located in the Swiss Alps (1) or outside the Swiss Alps (0).		
Comment			
Values	Binary	0,1	none
Reference			
Usage	**	This spatial filter must be applied, if modeling is done with the dataset **large**. A better alternative is to use the property **SKIPROX** as filter.	
Copyrights	© Skitourenguru		

The following image shows the extent of the spatial filter. With the traffic light colors the travel density calculated with a Kernel of 5000 m is shown: