

Visuelle Lawinenrisiko-Analyse

Ein Visual Analytics Dashboard basierend auf historischen Lawinenunfalldaten

Tamara Nyffeler

Fachhochschule Graubünden FHGR

Schweizerisches Institut für Informationswissenschaft

Chur, Schweiz

tamara.nyffeler@stud.fhgr.ch

Abstract—Diese Studie untersucht Lawinenunfälle im Schweizer Alpenraum, basierend auf 25 Jahren Daten des Schweizerischen Instituts für Schnee- und Lawinenforschung (SLF). Ein interaktives Dashboard, das tägliche Schnee- und Wetterdaten von automatischen Messstationen integriert, ermöglicht eine Risikoeinschätzung der aktuellen Lawinensituation in verschiedenen alpinen Regionen der Schweiz sowie einen Vergleich von Schnee- und Wettertrends vergangener Wintersaisons. Ziel der Untersuchung ist es, mit Hilfe verschiedener Visualisierungstechniken Muster und Zusammenhänge zwischen Lawinenunfällen und Wetterbedingungen aufzudecken.

Index Terms—Lawinenrisiko, SLF, IMIS-Messnetz, Schnee- und Lawinenforschung, Lawinenunfallereignisse, Schneetrends, Visual Analytics, Dashboard-Design

I. EINFÜHRUNG

Bergführer, Skitourengeher und Wintersport-Enthusiasten stehen oft vor der Herausforderung, Lawinensituationen im hochalpinen Raum schnell und präzise einschätzen zu können. Eine ansprechende und intuitive Visualisierung der relevanten Daten kann dabei eine wertvolle Unterstützung bieten. Aus diesem Grund wurde das Lawinen-Dashboard entwickelt; ein interaktives Visual-Analytics-Tool, das auf Python-Programmcode sowie den Bibliotheken Dash, Plotly, Pandas und Scikit-Learn basiert. Das Lawinen-Dashboard ist öffentlich zugänglich und kann über den URL snow-avalanche-analytics.onrender.com aufgerufen werden. Die zugrunde liegende Datenanalyse basiert auf aggregierten und gemittelten Tagesdaten, wobei Algorithmen wie K-Means [5] und Principal Component Analysis (PCA) [6] zum Einsatz kommen, um Muster und Risikogruppen zu identifizieren (Figure 1).

II. ANALYSIERTE DATEN

Die Datenbasis bildet eine Kombination aus statischen und täglich aktualisierten Live-Daten:

Statische Daten: Der erste Datensatz enthält Metadaten des Interkantonalen Mess- und Informationssystems (IMIS) [3]. Der zweite Datensatz umfasst historische Unfalldaten seit 1970/71 [10], einschliesslich Unfalltag, Startkoordinaten der Lawinenabgänge sowie Details zu verunfallten, geborgenen und verstorbenen Personen. Dieser Datensatz endet im Juni 2023 und umfasst rund 3'300 Unfallereignisse. Beide Datensätze sind als CSV-Datei verfügbar und dienen als Trainingsdaten für die Analyse-Algorithmen.

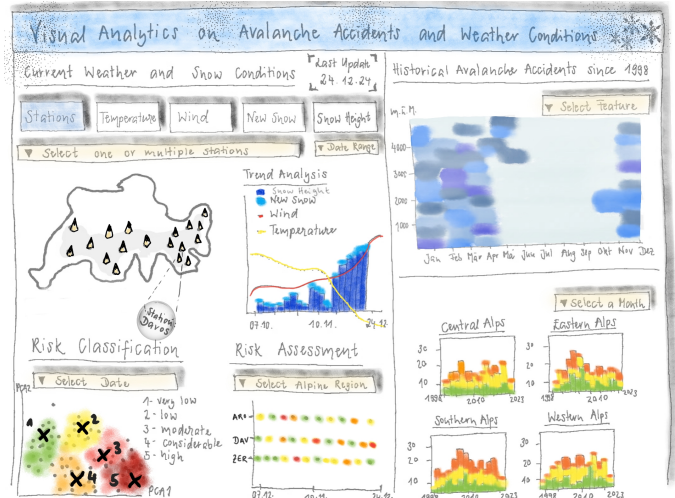


Fig. 1. Mockup des Visual-Analytics-Tool

Live-Daten: Die aktuellen Live-Daten werden vom SLF über eine Open-Data-API zur Verfügung gestellt [4], mit historischen Messwerten bis 1998. Die benötigten Daten werden in zwei separaten API-Calls bezogen. Der erste Abruf enthält Wetterdaten wie Windgeschwindigkeit (m/s), Luft- und Schneetemperatur (°C) aller IMIS-Messstationen, die halbstündlich aktualisiert werden. Der zweite Abruf liefert Schneedaten, darunter die durchschnittliche Schneehöhe und die Menge des Neuschnees (in cm) der letzten 24 Stunden. Die Schnittstelle gibt die Daten als JSON-Format aus, diese werden dann im Python-Code in ein Pandas-Dataframe konvertiert und als CSV gespeichert [7].

III. TRANSFORMATION DER TRAININGSDATEN

Das Datacleaning erfolgte zunächst in einer SQL-Datenbank, um einen Überblick über die Datengrösse und allgemeine Qualität zu erhalten. Anschliessend wurden die statischen Daten mit Python so transformiert, dass die Wetter- und Schneedaten den Unfalldaten zugeordnet werden konnten. Dieses Mapping erfolgte anhand der geografischen Distanz (Geospatial Mapping) unter Verwendung der Haversine-Formel [9]. Der Suchradius wurde auf 10 km festgelegt, sodass nur Stationen innerhalb dieses Umkreises berücksichtigt wurden. Für die ausgewählten Stationen wurden die Wetterparameter (Lufttemperatur, Windgeschwindigkeit, Schneehöhe,

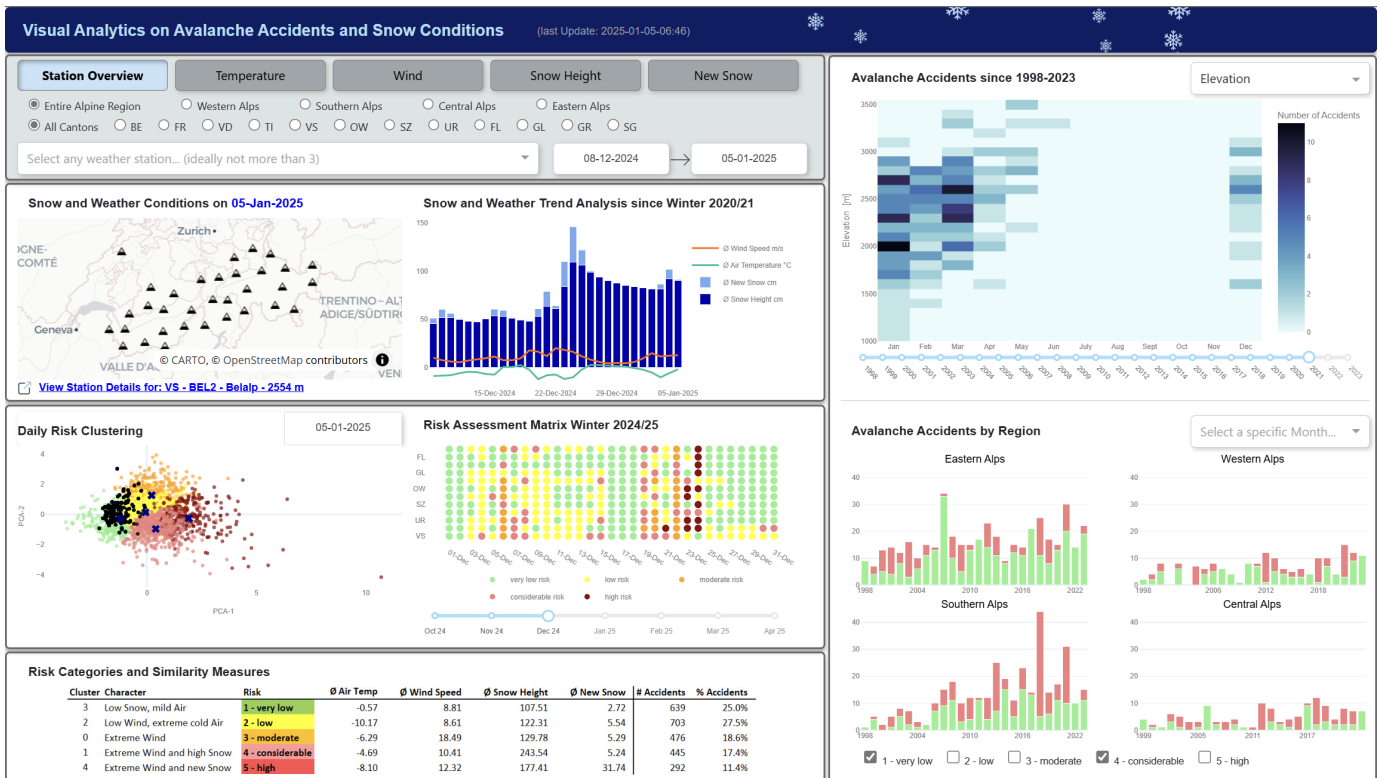


Fig. 2. Das Lawinen-Dashboard bietet verschiedene Ansichten: Im rechten Bereich ermöglicht eine Heatmap (oben) und ein gestapeltes Balkendiagramm (unten) die Exploration der Unfalldaten, die als Trainingsdaten dienen. Im linken Bereich werden verschiedene Visualisierungen der Live-Daten dargestellt, die durch einen Algorithmus nach Risikogruppen klassifiziert wurden. Die Ausprägungen dieser Kategorien sind unten links beschrieben.

Neuschnee) als Mittelwerte berechnet. In einem weiteren Schritt wurden die Unfalldaten anhand der Kantone in vier alpine Regionen unterteilt: Western-, Southern-, Central- und Eastern-Alps. Da die Wetteraufzeichnungen nur bis 1998 verfügbar sind, wurden auch die Unfalldaten auf diesen Zeitraum beschränkt. Der resultierende Datensatz, der Unfälle zwischen 1998 und 2023 umfasst, diente im Weiteren als Trainingsbasis für den Analyse-Algorithmus.

IV. AKTUALISIERUNG DER LIVE DATEN

Die tägliche Aktualisierung der Realtime-Daten ist über automatisierte Workflows in GitHub gesteuert und erfolgt in fünf Schritten:

- 1. Abruf neuer Wetterdaten:** API-Abruf und Fortschreibung in CSV-File SLF_daily_imis_measurements.
- 2. Abruf neuer Schneedaten:** API-Abruf und Fortschreibung in CSV-File SLF_daily_imis_snow.
- 3. Löschen von Duplikaten:** Bereinigung der beiden CSV-Files, um sicherzustellen, dass keine Duplikate bestehen.
- 4. Aggregation auf Tag:** Da die IMIS-Stationen halbstündlich Messwerte liefern, werden diese zu täglichen Aggregaten zusammengefasst.
- 5. Konsolidierung der Daten:** Die Wetter- und Schneedaten werden zu einem CSV-File zusammengefasst (SLF_daily_imis_all_live_data), um das weitere Datenhandling zu vereinfachen.

6. Risiko-Clustering: Der K-Means-Algorithmus wird verwendet, um die Trainingsdaten in fünf Risikogruppen zu unterteilen: very low, low, moderate, considerable und high. Die Live-Daten werden anschliessend den zuvor trainierten Clustern zugeordnet. Zur Visualisierung werden die Daten mittels PCA auf zwei Hauptkomponenten reduziert.

V. HYPOTHESEN

Für die Exploration der Daten wurden folgende vier Hypothesen aufgestellt:

- Hypothese 1: Korrelation Schneemenge und Lawinengefahr**
Grössere Schneemengen führen zu einem höheren Lawinenrisiko.
- Hypothese 2: Langfristige Schneeveränderung**
Über den Analysezeitraum hinweg lässt sich ein Trend in den jährlichen Schneemengen erkennen.
- Hypothese 3: Saisonale Verteilung**
Lawinenunfälle zeigen saisonale Muster, die zu Beginn, in der Mitte und am Ende des Winters variieren.
- Hypothese 4: Regionale Unterschiede**
Im Schweizer Alpenraum lassen sich regionale Unterschiede bei den Lawinenunfällen feststellen.

VI. VISUALISIERUNGSDESIGN

Zur Validierung der zuvor formulierten Hypothesen wurde ein Dashboard entwickelt, das sechs verschiedene Visualisierungskomponenten enthält, die im Folgenden erläutert werden (Figure 2):

- **Geografische Karte:** Die Geo-Map visualisiert die Standorte der IMIS-Stationen und zeigt die aktuellen Wetter- und Schneebedingungen als Overlay [2].
- **Kombiniertes Balken- und Liniendiagramm:** Diese Grafik zeigt die Schneehöhe, Windgeschwindigkeit und Lufttemperatur über die Zeit. Durch Auswahl eines spezifischen Wetterparameters wird die entsprechende Visualisierung als Liniendiagramm dargestellt, was die Trends der letzten fünf Jahre veranschaulicht.
- **Scatterplot:** In dieser Grafik wird die Klassifizierung der fünf Risikogruppen visualisiert. Die Trainingsdaten sind nach Risikogruppe von grün bis rot eingefärbt, während die täglichen Messwerte in schwarz dargestellt sind.
- **Punktematrix:** Diese Matrix liefert eine Übersicht der klassifizierten Risikogruppen, aufgeschlüsselt nach Region auf der Y-Achse und nach Tag auf der X-Achse.
- **Heatmap:** Diese Grafik stellt die Verteilung der Unfälle über das Kalenderjahr dar. Dunklere Felder weisen auf eine höhere Anzahl an Unfallereignissen hin. Die Parameter auf der Y-Achse können manuell ausgewählt werden.
- **Gestapeltes Balkendiagramm:** Diese Visualisierung stellt die Anzahl der Unfälle über alle Jahre hinweg dar, unterteilt in Small Multiples [2] für die vier alpinen Regionen. Die Risikokategorien sind farblich entsprechend der Codierung im Scatterplot und in der Punktematrix aufgebaut.

VII. INTERAKTIONSTECHNIKEN

Das Werkzeug integriert verschiedene interaktive Elemente, die zur Analyse und Visualisierung konzipiert sind. Die Benutzeroberfläche ist so gestaltet, dass sie eine möglichst effiziente und vielseitige Interaktion ermöglicht, wobei verschiedene Werkzeuge und Techniken zum Einsatz kommen.

- **Eingabekomponente:** Im oberen linken Bereich des User Interfaces befinden sich verschiedene Eingabekomponente wie fünf Schaltflächen, mehrere Radiobuttons und ein Dropdown-Menü (Figure 3). Dies ist der zentrale Ausgangspunkt für die Nutzerinteraktion und die Exploration der Live-Daten im linken Dashboard-Bereich. Die getroffene Auswahl wird auf die geografische Karte, die Trendanalyse und die Punktematrix angewandt. Die Trendanalyse kann zusätzlich über einen Kalender im Eingabefenster gesteuert werden, indem Start- und Enddatum für die Analyse festgelegt werden. Die 'Risk Assessment' Punktematrix verfügt ebenfalls noch über einen eigenen Schieberegler, mit dem der Monat gewechselt werden kann. Der Scatterplot 'Daily Risk Clustering' wird hingegen gänzlich über einen separaten Kalender gesteuert, der nur Tageswerte zulässt und keine Bereichsauswahl ermöglicht.

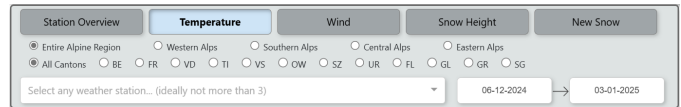


Fig. 3. Das Eingabefenster zur Exploration der Live-Daten

- **Hover-Funktion:** Alle Visualisierungen haben eine Hover-Funktion, die zusätzliche Informationen zu den Datenpunkten anzeigt (Figure 4). Zudem ermöglicht die Zoom-In-Funktion eine genauere Auswahl und Detailansicht überlappender Daten.

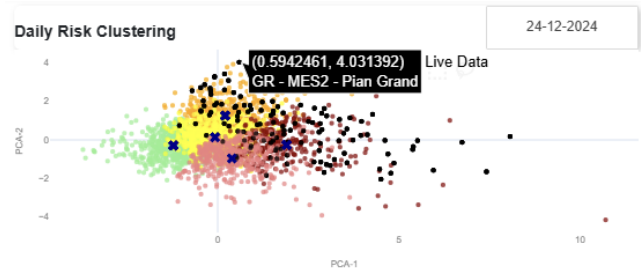


Fig. 4. Beispiel der Hover-Funktion im Scatterplot

- **Interaktive Links:** Die geografische Karte mit den IMIS-Stationen bietet interaktive Links, die den direkten Zugriff auf das Datenblatt der jeweiligen Messstation ermöglichen (Figure 5). Der Link unterhalb der Karte wird dynamisch aktualisiert, sobald eine andere Station ausgewählt wird. Das Datenblatt öffnet sich in einem neuen Tab.



Fig. 5. Absprung auf URL zur Detailinformation der IMIS-Station

- **Dropdown-Menü:** Das Visual-Analytics-Tool hat drei Dropdown-Menüs. Das erste Menü befindet sich im Eingabebereich und ermöglicht die Auswahl spezifischer IMIS-Stationen. Es wird empfohlen, nicht mehr als drei Stationen gleichzeitig auszuwählen, um eine Überladung der Informationen zu vermeiden (Visual-Clutter Effekt [2]). Die beiden anderen Dropdown-Menüs befinden sich im rechten Dashboard-Bereich. Eines davon wird verwendet, um in der Heatmap die Y-Achse zu steuern (Figure 6) und das andere, um im gestapelten

Balkendiagramm einen bestimmten Monat der Unfalldaten anzuzeigen.

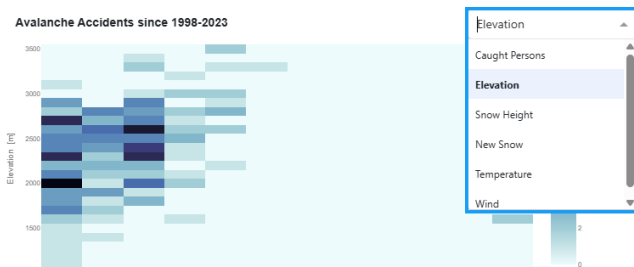


Fig. 6. Auswahl der Y-Achse in der Heatmap der Unfalldaten

- **Check-Boxen:** Da ein gestapeltes Balkendiagramm unübersichtlich wird, wenn zu viele Balken übereinander liegen [8], können einzelne Risikokategorien über Check-Boxen ein- und ausgeblendet werden (Figure 7). Die beste Vergleichbarkeit wird erreicht, wenn maximal zwei Kategorien gleichzeitig aktiviert sind.

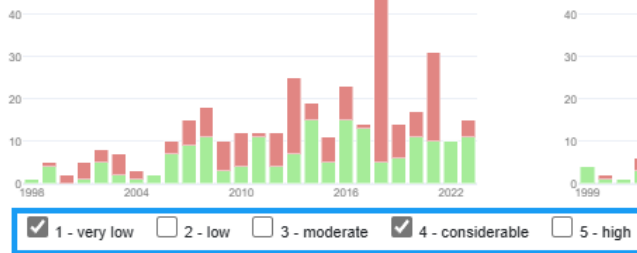


Fig. 7. Check-Boxen zur Auswahl der Risikokategorien

VIII. PERFORMANZANALYSE

Die Performanzanalyse fokussiert sich auf zwei zentrale Aspekte: die Ladezeit und die Skalierbarkeit. Beide sind entscheidend für die Nutzererfahrung und die Effizienz des Dashboards [1].

- **Ladezeit:** Wie in Kapitel IV beschrieben, wurde die Datenaktualisierung und die algorithmische Analyse in automatisierte Prozesse ausgelagert, die täglich zwischen 07:00 und 08:00 Uhr ablaufen und weniger als eine Minute in Anspruch nehmen. Beim Aufruf des Dashboards müssen daher keine rechenintensiven Tasks durchgeführt werden – es werden lediglich die vorbereiteten CSV-Files geladen und die Daten visualisiert, was weniger als eine Sekunde dauert. Beim ersten Aufruf nach längerer Inaktivität kann die Ladezeit dennoch bis zu einer Minute dauern, da der Host zuerst aktiviert werden muss.
- **Skalierbarkeit:** Es ist hervorzuheben, dass das Lawinen-Dashboard primär für die Nutzung auf PC-Monitoren konzipiert ist und auf mobilen Geräten mit kleinerer Bildschirm diagonale keine optimale Darstellung bietet. Das Tool weist im aktuellen Stand kein responsives Design auf.

IX. ERKENNTNISSE

Der Erkenntnistransfer basiert auf einer umfangreichen Datenbasis, die mit Hilfe des entwickelten Lawinen-Dashboards analysiert wurde, wobei der Fokus auf den in Kapitel V aufgestellten Hypothesen liegt:

- **Korrelation Schneemenge und Lawinengefahr:** Die Punktematrix veranschaulicht den Zusammenhang zwischen starkem Schneefall und einem erhöhten Lawinenrisiko. So führte beispielsweise der intensive Schneefall an Weihnachten 2024 in fast allen Kantonen zu einem deutlichen Anstieg des Lawinenrisikos. Ein Blick auf die historischen Daten zeigt einen ähnlichen Zusammenhang: Im schneereichen Winter 2021 wurden die meisten Lawinenunfälle im Analysezeitraum verzeichnet.
- **Langfristige Schneeänderung:** In der Trendanalyse wird ein deutlicher Kontrast zwischen den schneereichen Wintern 2021 und 2024 und den schneeärmeren Wintern 2022 und 2023 sichtbar. Diese Unterschiede verdeutlichen die Variabilität der Schneemengen im betrachteten Zeitraum.
- **Saisonale Verteilung:** Die Analyse der Heatmap legt nahe, dass sich die Unfalldichte zum Ende der Wintersaison verstärkt in höhere Lagen bewegt. Dieser Trend lässt sich darauf zurückführen, dass in tieferen Lagen Schneebedingungen zunehmend unzureichend werden, wodurch sportliche Aktivitäten verstärkt in höher gelegene Regionen verlagert werden.
- **Regionale Unterschiede:** Die Regionen mit der höchsten Unfalldichte korrelieren oft mit beliebten Wintersportgebieten, wie beispielsweise den Eastern-Alps (Kanton Graubünden) und den Southern-Alps (Kanton Wallis), was sich auch in der Unfalldatenbank widerspiegelt.

X. FAZIT UND WEITERFÜHRENDE ARBEIT

Das Werkzeug bietet vielfältige Interaktionsmöglichkeiten, um den vorhandenen Datensatz zu explorieren. Die aktuellen Visualisierungen stellen jedoch nur einen Anfang dar und bieten Raum für weitere Ausgestaltung und Verfeinerung. Es ist wichtig zu beachten, dass sich die Analysen auf vier wesentliche Parameter beschränken: Windgeschwindigkeit, Lufttemperatur, Schneehöhe und Neuschneemenge. Die Natur der Analysen ist explorativ, weshalb das Dashboard nicht als alleinige Entscheidungsgrundlage für sicherheitskritische Situationen genutzt werden sollte.

REFERENCES

- [1] M. Burch. *Dashboard Design*. River Publishers Series in Computing and Information Science and Technology Series. River Publishers.
- [2] M. Burch. Skalierbare visualisierungstechniken.
- [3] W. I. for Snow and A. R. SLF. Description of automated stations.
- [4] W. I. for Snow and A. R. SLF. Slf measurement api.
- [5] A. M. Grellmann. Einführung in den k-means-algorithmus und seine implementierung in python.
- [6] N. Lang. Principal component analysis data basecamp.
- [7] K. Mehreen. How to convert JSON data into a DataFrame with pandas.
- [8] J. A. Schwabish. *Better data visualizations: a guide for scholars, researchers, and wonks*. Columbia University Press.
- [9] P. Teacher. The haversine formula for geospatial distances.
- [10] B. Zweifel. Avalanche accidents in switzerland since 1970/71.