

---

# **Digitale Erfassung und Inventarisierung von gravitativen Massen- bewegungen entlang von Bundesfernstraßen**

---

Fachveröffentlichung der  
Bundesanstalt für  
Straßenwesen

---

# Digitale Erfassung und Inventarisierung von gravitativen Massenbewegungen entlang von Bundesfernstraßen

---

von

Michael Steffen, Thomas Hertwig, Mandy Schipek, Bernd Torchalla  
Beak Consultants GmbH, Freiberg

---

## Impressum

Fachveröffentlichung zu Forschungsprojekt: 89.0337

Digitale Erfassung und Inventarisierung von gravitativen Massenbewegungen entlang von  
Bundesfernstraßen

Fachbetreuung:  
Anne-Farina Lohrengel

Referat:  
Nachhaltiges Bauen und Erhalten von Straßen

Herausgeber:  
Bundesanstalt für Straßenwesen  
Brüderstraße 53, D-51427 Bergisch Gladbach  
Telefon: (0 22 04) 43 - 0

<https://doi.org/10.60850/FV-S-89.0337>

Bergisch Gladbach, Juli 2024

Zu diesem Forschungsprojekt werden nur die Kurzfassung und der Kurzbericht veröffentlicht.  
Die Langfassung des Schlussberichts kann auf Anfrage an [verlag@bast.de](mailto:verlag@bast.de) zur Verfügung ge-  
stellt werden.

Es wird darauf hingewiesen, dass die unter dem Namen der Verfasser veröffentlichten Berichte  
nicht in jedem Fall die Ansicht des Herausgebers wiedergeben. Nachdruck und photomechani-  
sche Wiedergabe, auch auszugsweise, nur mit Genehmigung der Bundesanstalt für Straßenwe-  
sen, Stabsstelle Presse und Kommunikation.

## **Kurzfassung · Abstract**

### **Digitale Erfassung und Inventarisierung von gravitativen Massenbewegungen entlang von Bundesfernstraßen**

Ziel des Forschungsvorhabens war die Entwicklung eines Verfahrens, welches zurückliegende gravitative Massenbewegungen auf Basis hochauflösender digitaler Höhendaten erfasst und klassifiziert sowie die einzelnen Prozesse in Form geeigneter Werkzeuge für das Geoinformationssystem Esri ArcGIS bereitstellt. Für die Erfüllung dieses Forschungsziels wurde zunächst eine Literaturanalyse durchgeführt und darauf aufbauend die grundlegende Vorgehensweise abgeleitet.

Durch die Berechnung von morphologischen Parametern auf Basis der digitalen Höhenmodelle wurden die charakteristischen Oberflächeneigenschaften bekannter Massenbewegungsprozesse ermittelt und mit deren Hilfe unter Einsatz von Machine-Learning-Algorithmen und einer objektbasierten Herangehensweise die generelle Machbarkeit des Verfahrens getestet sowie verschiedene Trainingsszenarien erstellt und der Modellaufbau optimiert.

Zur Anwendung kamen die Verfahren der Künstlichen Neuronalen Netze sowie der Zufallswälder. Die Durchführung erfolgte bundeslandweit für Nordrhein-Westfalen und Sachsen in Versionen mit 2 m und 5 m Pixelgröße. Nach Abschluss der Berechnungen wurden die Unterschiede der Modellergebnisse hinsichtlich der Machine-Learning-Verfahren sowie der Auflösung herausgearbeitet und beurteilt, wobei teils signifikante Unterschiede erkennbar wurden. Dies spiegelt sich vor allem in Lage und Anzahl der ermittelten Objekte wider, bezogen auf den kleinmaßstäblichen Bereich werden jedoch prinzipiell die gleichen Regionen ausgewiesen.

Weiterhin erfolgte eine Attributierung der ermittelten Flächen bezüglich Größe, Position, Hangneigung, Geologie und Landnutzungsart sowie die Kennzeichnung potentiell stark betroffener Straßenabschnitte. Der darauf aufbauende Abgleich auf Plausibilität ergab, dass sich im Durchschnitt ca. ein Fünftel der Objekte in anthropogenen Bereichen befindet. Unter Einbeziehung geologischer Gegebenheiten konnten weiterhin potentielle aktive Sturzgebiete identifiziert werden.

### **Digital site identification and inventory of gravitational mass movements along the national trunk road network**

The aim of this research project was the development of a method that identifies and classifies post-event gravitational mass movements on the basis of high-resolution digital elevation data and furthermore provides appropriate tools for the geographic information system Esri ArcGIS. For the fulfilment of this research objective, a literature analysis was first carried out. Based on this, a primary procedure was inferred.

By calculating morphological parameters on the basis of the digital elevation models, characteristic surface properties of known gravitational mass movements were determined and in combination with machine learning algorithms as well as an object-based approach, the general methodical feasibility was tested and various training scenarios were created. Thus, the models were stepwise improved.

For the application, artificial neural networks and random forests were used. The implementation was done for the areas of the federal states of North Rhine-Westphalia and Saxony in versions with 2 m and 5 m pixel size. After completion of the calculations, the model results were analysed and evaluated with regard to the machine learning methods and their respective resolution. At this, the results showed up significant differences in part, which becomes apparent by observing both the position as well as the number of the identified objects. Though, relating to a small-scale range, the detected areas are basically located in the same regions.

Furthermore, an attribution of the identified areas regarding size, position, slope, geology and land use was carried out and potentially heavy affected road sections were distinguished. Based on this information, a test of plausibility revealed that, on average, about one fifth of the identified objects are located in anthropogenic affected areas. Taking into account geological conditions, potentially active rockfall areas could be identified.

## Kurzbericht

# Digitale Erfassung und Inventarisierung von gravitativen Massenbewegungen entlang von Bundesfernstraßen

## 1 Ausgangslage und Zielstellung

Gravitative Massenbewegungen sind in Deutschland weit verbreitete Gefahren und können z.T. erhebliche Schäden an der vorhandenen Verkehrsinfrastruktur hervorrufen. Neben Hang- und Böschungsrutschungen, Felsstürzen und Geröllabgängen zählen auch Fließprozesse wie Schutt- und Schlammströme in diese Kategorie der meist klimatisch gesteuerten Naturgefahren. In der Nähe von Verkehrswegen können sie nicht nur zu baulichen Schäden, sondern auch zu starken Beeinträchtigungen des Verkehrs und damit verbunden ebenfalls zu weiteren volkswirtschaftlichen Kosten führen.

Das Ziel des vorliegenden Forschungsvorhabens besteht deshalb darin, ein Verfahren zu entwickeln, welches zurückliegende gravitative Massenbewegungen erfasst und klassifiziert. Die Identifizierung und Analyse der kartierten Ereignisse können anschließend als Basis für weitere Ansätze zum Monitoring gefährdeter Hänge sowie für die Früherkennung potentieller Schadensfälle dienen. Als Datengrundlage sollen die auf Länderebene verfügbaren digitalen Geländemodelle mit 1 m Auflösung verwendet werden. Die Methodik ist dabei so zu entwickeln und aufzubereiten, dass der Auftraggeber nach Abschluss des Projektes selbstständig in der Lage ist, die Arbeiten fortzuführen und neue Gebiete hinzuzufügen sowie bereits vorhandene Bereiche zu aktualisieren. Die hierfür zu erstellenden Werkzeuge sollen in einer ArcGIS Desktop-Umgebung lauffähig sein.

## 2 Kenntnisstandsanalyse

Die Analyse der Literatur aus den vergangenen zwei Jahrzehnten ergab eine Häufung der Veröffentlichungen in den Zeitschriften Computers & Geosciences, Geomatics, Geomorphology, Natural Hazards and Risk sowie Natural Hazards. National hingegen ist das Rechercheergebnis im Wesentlichen auf wissenschaftliche Qualifizierungsarbeiten, technische Berichte oder auch Tagungsschriften begrenzt. Die Veröffentlichungen wurden zunächst thematisch grob in die drei Kategorien Rutschprozesse, Sturzprozesse und Fließprozesse eingeteilt, wobei sich ein klarer Fokus auf Rutschungsprozesse zeigt. Publikationen zu Sturzprozessen sind vor allem geotechnischer Natur und in deutlich geringerer Anzahl vertreten. Der Teilbereich Bodenfließen scheint derzeit unter dem Aspekt der Erkennung von Massenbewegungen praktisch keine Beachtung zu finden. Die überwiegende Mehrheit der Veröffentlichungen beschäftigt sich zudem mit der Erstellung von Georisikokarten für potentielle Hangrutschungen, erst seit etwa fünf bis sechs Jahren werden zunehmend Studien mit objektorientierten Methoden zur Identifizierung vergangener Ereignisse durchgeführt, wobei die Erkennung selbst mit Hilfe überwachender Lernalgorithmen realisiert wird. Hinsichtlich der Validierungsergebnisse in den gesichteten Studien konnten nur selten mehr als 70 % der Flächen korrekt als Ereignis klassifiziert werden. Die visuelle Kontrolle der Ergebnisse spielt nach wie vor eine tragende Rolle zur Einschätzung der Modellqualität.

## 3 Datenbeschaffung- und Datenaufbereitung

Zur Bearbeitung des Projektes standen die folgenden Ausgangsdaten zur Verfügung:

- das Bundesfernstraßennetz aus dem Bundesinformationssystem Straße, © BISStra/BASSt (2017),
- die geologische Übersichtskarte GÜK200, © GeoBasis-DE/BKG (2013),
- das digitale Landschaftsmodell ATKIS Basis-DLM, © GeoBasis-DE/BKG (2018),
- die hochauflösenden digitalen Höhenmodelle der Bundesländer Nordrhein-Westfalen (1 m, © Land NRW (2018)), Sachsen (2 m, © Land Sachsen (2018)) und Thüringen (1 m, © Land Thüringen (2018)),
- Daten aus Ereigniskatastern.

Die Ereignisdaten wurden zunächst in die Hauptprozesstypen Fließprozesse, Sturzprozesse sowie Rutschungsprozesse klassifiziert, inhaltlich auf Plausibilität geprüft sowie deren Lagebezug zum Höhenmodell validiert. Dabei konnten Fließprozesse mit ca. 68 % sowie Sturzprozesse mit ca. 58 % als prinzipiell plausibel in Bezug zu ihrer Position eingestuft werden, einzelne Ereignisse werden jedoch durch das Relief

nicht abgebildet. Eine Kartierung bzw. Erkennung einzelner Steinschläge, Felsabbrüche oder ähnlicher Prozesse ist auf Grundlage der vorhandenen Daten deshalb nicht möglich. Bezüglich den Rutschungsprozessen konnten nur ca. 15 % der Ereignisdaten durch den Abgleich mit dem Höhenmodell bestätigt werden. Der geringe Prozentsatz ist dabei weitestgehend durch die Größe, das Alter und die Erosionsbeständigkeit sowie der anthropogenen Beeinflussung bzw. vom Beräumungsgrad abhängig.

## 4 Konzept

### 4.1 Maschinelles Lernen und Software

Entsprechend den Ergebnissen der Literaturanalyse wird sich bei der Erstellung der Methodik ebenfalls auf den Ansatz mittels überwachender Lernalgorithmen konzentriert. Weiterhin wird eine objektorientierte Herangehensweise favorisiert. Für die Entwicklung eines Algorithmus bzw. einer Methodik zur Erkennung verschiedener Massenbewegungsprozesse werden dabei hauptsächlich die Verfahren der Künstlichen Neuronalen Netze und Zufallswälder in Betracht gezogen. Softwareseitig stehen dafür die Anwendungen Esri ArcGIS Desktop 10.5.1 (32-Bit), Beak advangeo® (32-Bit) sowie das Python-3-Paket Scikit-learn (64-Bit) zur Verfügung. Die Datenebenen, welche für das Maschinelle Lernen notwendig sind, werden in Form von Reliefparametern durch die Ableitung aus den digitalen Höhenmodellen bereitgestellt.

### 4.2 Reliefparameter

Die in der Literatur genutzten Reliefparameter begrenzen sich meist auf Modelle, welche zum Zweck der Georisikoanalyse erstellt werden – nicht aber zur Erkennung von Strukturen – weshalb vor allem Landnutzungsdaten und hydrologische Parameter sowie Informationen zum Boden bzw. der vorhandenen Geologie Verwendung finden. Die relevanten Parameter zur Beschreibung der Oberflächenbeschaffenheit in derartigen Modellen sind häufig nur die Rauigkeits- und Positionsindizes, Hangneigung, Hangneigungsrichtung und ferner die Plan- und Profilkümmungen.

Da diese Parameter allein nicht genug Spielraum für die Modellberechnungen bieten, um die Oberflächen und Segmente zu differenzieren, werden darüber hinaus weitere Reliefableitungen aus verschiedenen Literaturquellen zur Beschreibung der Oberflächenmorphologie einbezogen. Insgesamt stehen 25 Parameter zur Verfügung.

### 4.3 Pixelbasierter vs. objektbasierter Ansatz

Jedes Pixel innerhalb eines Rasters besitzt einen spezifischen Wert (z.B. Farbwert, Magnitude eines Krümmungsparameters, Hangneigung, Höhenwert des digitalen Höhenmodells etc.), welcher zunächst von den Pixelwerten der ihn umgebenden Punkte weitestgehend unabhängig ist, so dass die Pixel praktisch keine Kenntnis voneinander besitzen. Da es sich jedoch im Speziellen bei den Rutschungsereignissen um Flächen handelt, in denen die Pixelwerte auf Grund der gemeinsamen Umlagerung von Material praktisch eine stärkere Abhängigkeit voneinander erlangen, ist die Herangehensweise auf Basis einzelner unabhängiger Bildpunkte nicht geeignet, weshalb auf Basis eines ausgewählten Reliefparameters im Vorfeld eine Segmentierung vorgenommen wird. Diese Objektbildung hat zum Ziel, einzelne Pixel zu Bereichen mit ähnlichen Eigenschaften zusammenzufassen sowie die Gebiete mit verschiedenen Oberflächenausprägungen möglichst präzise voneinander abzugrenzen.

Zur Nutzung in einem Machine-Learning-Ansatz werden die auf diesem Wege generierten Objekte zusätzlich mit den Werten der aus den digitalen Höhenmodellen abgeleiteten Reliefparameter belegt. Dies wird realisiert, indem für jeden Oberflächenparameter sowohl Mittelwert, Standardabweichung als auch Wertebereich pro Objekt kalkuliert wird, wodurch sich zudem die Anzahl der Eingangsebenen für die Anwendung mit den Künstlichen Neuronalen Netzen bzw. Zufallswäldern verdreifacht.

## 5 Training und Modellaufbau

### 5.1 Herangehensweise

Das Ergebnis der Identifizierung gravitativer Massenbewegungen auf Basis ihrer morphologischen Eigenschaften ist von einer Vielzahl an Einflussgrößen abhängig:

- der Aufbereitung des Höhenmodells: Filterung/Glättung,
- der verwendeten Auflösung der Ausgangsdaten,
- der einbezogenen Reliefparameter,
- dem Ergebnis der Segmentierung und Objektbildung,
- der Anzahl und Qualität kartierter Massenbewegungen (Lernflächen) sowie
- dem Verfahren des Maschinellen Lernens.

Entsprechend dieser Liste an Faktoren existiert auch eine Fülle von Kombinationsmöglichkeiten für den Modellaufbau. Zur Beurteilung der o.g. Punkte wurde deshalb eine Vielzahl an Testmodellen mit jeweils veränderten Faktoren berechnet und anschließend verglichen. Für die Auswertung der Ergebnisse wird sowohl der visuelle Eindruck als auch das ermittelte Verhältnis von richtig und falsch vorhergesagten Flächen genutzt.

Eine wichtige Erkenntnis ist hierbei, dass durch Verwendung des 1-m-Höhenmodells gegenüber dem 2-m-Höhenmodell keine signifikante Verbesserung der Ergebnisse beobachtet werden konnte. Die Nutzung des Modells mit 2 m Zellgröße führt somit zu einer erheblichen Reduzierung der Berechnungszeiten als auch des erforderlichen Speicherplatzes bzw. des Datenvolumens. Das höchste Potential zur Verbesserung der Trainingsergebnisse scheint dabei jedoch in den verfügbaren Trainingsflächen zu liegen. Auf Grund des geringen Umfangs vorhandener Lernflächen ist davon auszugehen, dass die Modelle bei Erweiterung der Datenbasis von repräsentativen und im Höhenmodell auch deutlich nachvollziehbaren Massenbewegungen stark verbessert werden können. Weiterhin konnte gezeigt werden, dass sowohl die Wahl als auch die Anzahl der Reliefparameter ebenfalls wichtige Faktoren hinsichtlich der Modellqualität sind. Die Verwendung von wenigen Reliefparametern (z.B. ArcGIS-Funktion: maximal 2 Parameter) führte zu keinem zufriedenstellenden Ergebnis.

### 5.2 Training der Modelle zur Anwendung

#### 5.2.1 Vorbemerkungen

Auf Grund des enormen Datenvolumens wurde zum Training mittels der Künstlichen Neuronalen Netze bzw. der Zufallswälder ein Trainingsmosaik erstellt. Hierfür wurden die Ereignisflächen in jede Richtung um 500 m gepuffert, ausgeschnitten und anschließend zu einem Mosaik zusammengefügt. Das erzeugte Trainingsmosaik besteht entsprechend der Anzahl verwendbarer Ereignisse aus 4 x 4 Kacheln und besitzt eine Ausdehnung von 2.000 x 2.000 Pixeln (2 m Zellgröße). Auf Grund technischer Begrenzungen der 32-Bit-Software können bei dieser Pixelanzahl nur ca. 11 Reliefparameter einbezogen werden. Da jedoch der Bestand an verwendbaren Trainingsflächen noch nicht sehr umfangreich ist, muss für die zukünftige Verwendung der Methodik auch die Funktionalität mit weitaus mehr Trainingsflächen gegeben sein. Aus diesem Grund wurden zusätzlich identische Modelle mit 5 m Zellgröße berechnet, was die mögliche verwendbare Kachelanzahl bereits auf ca. 150 anhebt. Die Umrechnung von 2 m auf 5 m wurde dabei mit den vorbereiteten Datensätzen der 2-m-Modelle durchgeführt.

Anschließend erfolgte die Erstellung der finalen Trainingsmodelle zur landesweiten Anwendung auf die Bundesländer Nordrhein-Westfalen und Sachsen. Zur Ermittlung des optimalen Modellaufbaus wurden hierfür im Vorfeld verschiedene Testszenarien erstellt, berechnet und ausgewertet. Da die Wichtung der eingesetzten Parameter bei jedem Training neu initiiert wird, wurden zum Einsatz der Künstlichen Neuronalen Netze je Auflösungsversion 5 identische Modelle generiert.

## 5.2.2 Ergebnisse

Der Vergleich der Trainingsergebnisse zwischen den Versionen mit 2 m und 5 m Zellgröße brachte keine grundlegenden Unterschiede zum Vorschein. Anders verhält es sich beim Abgleich der Ergebnisse der Künstlichen Neuronalen Netze und der Zufallswälder, bei denen teils deutliche Unterschiede beobachtet werden konnten. Besonders gravierend äußert sich dies bei Betrachtung der relativen Wichtung einzelner Reliefparameter, die sowohl bereits bei den Modellen der Künstlichen Neuronalen Netze sichtbar schwankt, als auch zwischen beiden Verfahren deutliche Unterschiede hervorbringt. Damit verbunden weichen auch das visuelle Erscheinungsbild sowie die Rate der korrekt bzw. falsch klassifizierten Flächen mit scheinbaren Vorteilen für das Verfahren der Zufallswälder teils merklich voneinander ab. Gründe hierfür können sein:

- a) Es stehen nicht genügend Trainingsflächen zur Verfügung.
- b) Die Zusammenfassung der Prozesse zum Training ist nicht förderlich.
- c) Die Morphologie wird durch die verwendeten Oberflächenparameter noch nicht gut genug beschrieben.
- d) Die Zufallswälder sind möglicherweise das geeignetere Verfahren zur Klassifikation.

Dennoch wird im Training der Modelle mit den Künstlichen Neuronalen Netzen zwischen 60 % und 70 % der Ereignisfläche korrekt klassifiziert. Bei den Zufallswäldern liegt die Erkennungsrate des Trainings für die richtig ausgewiesene Fläche bei fast 100 %, was jedoch auch als Zeichen für eine Überanpassung des Modells gedeutet werden kann.

# 6 Anwendung

## 6.1 Vorbemerkungen

Zur Anwendung der Künstlichen Neuronalen Netze wurden für jede Auflösungsversion die beiden Trainingsmodelle mit den geringsten Fehlern ausgewählt und die Berechnung für die gesamte Landesfläche durchgeführt. Für die Ergebnisse im Wertebereich zwischen 0 (Ereignis unwahrscheinlich) und 1 (Ereignis wahrscheinlich) wurde nachfolgend der Durchschnitt gebildet und die Klassifikation in „Ereignis/kein Ereignis“ auf Basis eines empirisch ermittelten Grenzwertes vorgenommen. Da das verwendete Python-3-Paket Scikit-learn zur Anwendung der Zufallswälder bereits eine Möglichkeit für die direkte Klassifikation bietet, wurde nur jeweils ein Modell je Auflösungsversion berechnet, so dass die Anwendung deutlich effektiver gestaltet werden konnte.

## 6.2 Ergebnisse

Die landesweite Anwendung der 2-m-Versionen mit den Künstlichen Neuronalen Netzen liegt mit Rechenzeiten von ca. 140 h für Sachsen bzw. ca. 410 h für Nordrhein-Westfalen weit über einem praktikablen Maß. Das Verfahren der Zufallswälder benötigt hingegen mit 20 h für Sachsen bzw. 33 h für Nordrhein-Westfalen vergleichsweise wenig Zeit. Bei den niedriger aufgelösten 5-m-Versionen rückt die Berechnungszeit der Künstlichen Neuronalen Netze für Nordrhein-Westfalen mit 18 h erheblich näher an die ca. 12 h der Zufallswälder heran. Beide Verfahren erreichen hier eine akzeptable Dauer.

Weiterhin konnte festgestellt werden, dass einzelne Pixel, kleinere Pixelgruppen und auch zum Teil anthropogene Strukturen, wie z.B. schmale Straßen oder Wege, als Erosionsereignisse klassifiziert werden, so dass eine zusätzliche Bereinigung der primären Ergebnisse notwendig wird.

Abhängig von Verfahren und Auflösung sind zudem deutliche Differenzen sowohl bezüglich der Anzahl als auch der Lage bei den identifizierten Flächen erkennbar. Die Betrachtung einzelner, spezieller Ereignisse ist deshalb nicht zu empfehlen, was besonders für relativ kleine Flächen, etwa unter einem halben Hektar gilt. Prinzipiell lassen die Ergebnisse jedoch einen Vergleich in kleineren Maßstäben zu, da bei der Erkennung im Wesentlichen Flächen in den gleichen Regionen ausgewiesen werden. Eine Verbesserung der Ergebnisse sollte jedoch mit weiteren zusätzlichen Ereignisdaten erzielt werden können.

### 6.3 Validierung

Die Attributierung und Auswertung der Landnutzungstypen ergab, dass sich ca. 15 % bis 20 % der klassifizierten Ereignisse innerhalb mehr oder weniger anthropogen beeinflusster Gebiete befinden, was eine deutlich erhöhte Gefahr für Fehlklassifikationen mit sich bringt. Demgegenüber stehen jedoch zwischen ca. 65 % und 80 % Flächen, die sich vorrangig in Waldgebieten befinden, so dass hier von einer geringeren Fehlklassifikationsrate ausgegangen werden kann.

Der Abgleich der Klassifikationsergebnisse mit der mittleren Hangneigung brachte besonders im Bundesland Sachsen eine relativ hohe Anzahl von Flächen in Klassen mit höheren Neigungswerten hervor. Dieser Effekt lässt sich damit begründen, dass in Sachsen, z.B. im Elbsandsteingebirge, viele steile Hänge im Festgestein existieren, welche auf Grund ihrer ausgeprägten Morphologie als Rutschungsprozess ausgewiesen wurden. In Bezug zur vorliegenden Aufgabenstellung stellt dies zwar eine Fehlklassifikation dar, kann jedoch unter zusätzlicher Einbeziehung der Geologie auf Gebiete mit aktiven Sturzprozessen hindeuten, auch wenn hiermit keine einzelnen Stürze erkannt werden können.

## 7 Schlussfolgerungen und Ausblick

Obwohl die Erstellung, Aktualisierung und Verbesserung der digitalen Höhenmodelle von den zuständigen Behörden kontinuierlich vorangetrieben wird, sind Aktualisierungsintervalle von mehreren Jahren keine Seltenheit, weshalb von der Entwicklung auf Basis einer multitemporalen Datengrundlage abgesehen wurde. Dieser Ansatz wäre dabei prinzipiell gut geeignet, um Veränderungen in den Höhenmodellen zu kartieren, jedoch kommen für diese Herangehensweise nur Daten mit entsprechend kurzen Aktualisierungsintervallen, wie z.B. hochauflösende Satellitenbilddaufnahmen, in Frage. Die Durchführung einer gezielten Analyse von bisher zu diesem Thema veröffentlichten Studien und grundsätzlich auch zur Vielfalt und Verfügbarkeit dieser Daten sowie darauf aufbauend die Entwicklung einer alternativen Methodik für die Beobachtung einzelner bekannter Ereignisse bzw. zu flächendeckenden Monitoringzwecken ist daher empfehlenswert.

Hinsichtlich der in diesem Forschungsvorhaben erhaltenen Ergebnisse erwies sich besonders die Verfügbarkeit geeigneter Lernflächen als unzureichend und dementsprechend als einschränkender Faktor. Bezüglich der technischen Umsetzung sind auf Grund der erheblich besseren Rechenzeiten sowie der Möglichkeit, auch die landesweite Anwendung mit 2 m Zellgröße in einem überschaubaren zeitlichen Rahmen durchzuführen, die Zufallswälder dem Künstlichen Neuronalen Netz gegenüber etwas im Vorteil. Die Nutzung der höchsten verfügbaren Auflösung von 1 m Zellgröße erscheint nicht praktikabel, da das Aufwand-Nutzen-Verhältnis gegenüber 2 m bzw. 5 m Pixelgröße unverhältnismäßig stark zunimmt. Da sich in den Anwendungsergebnissen der Kartierung im Detail nicht nur signifikante Unterschiede zwischen den Verfahren, sondern im Gegensatz zu den Trainingsergebnissen auch zwischen den Auflösungsversionen zeigten, ist eine valide Einschätzung einzelner Flächen nur durch die händische Überprüfung des digitalen Höhenmodells oder alternativ einer Feldbegehung möglich.

Weiterhin konnte aufgezeigt werden, dass sowohl die Modelle der Künstlichen Neuronalen Netze als auch der Zufallswälder grundsätzlich die gleichen Regionen ausweisen, auch wenn Ausdehnung, Lage und Anzahl der Flächen im Detail nicht direkt vergleichbar sind. Darüber hinaus können unter Beachtung der Geologie großflächig Gebiete eingestuft werden, welche nicht als Rutschungs- oder Bodenfließprozesse, sondern vielmehr als potentielle Bereiche mit aktiven Sturzprozessen zu interpretieren sind. Besonders deutlich wird dies im Bundesland Sachsen, z.B. im Elbsandsteingebirge.

Abschließend kann festgehalten werden, dass sich die grundsätzliche Vorgehensweise und Methodik zur initialen Erkennung von Massenbewegungen auf Basis des hochauflösenden digitalen Höhenmodells trotz einiger Fehlklassifikationen, wie z.B. in anthropogenen Bereichen, als anwendbar erwiesen hat. Basierend auf der vorhandenen Datengrundlage sowie den Unterschieden in den Anwendungsergebnissen beider verwendeter Verfahren lassen sich jedoch keine genaueren Aussagen zur Position einzelner Flächen und deren qualitativer Beurteilung treffen. Eine Anwendung des Algorithmus zu Monitoringzwecken kann dabei auf Grund der verhältnismäßig großen Aktualisierungsintervalle und Verfügbarkeiten der digitalen Höhenmodelle zum aktuellen Zeitpunkt ausgeschlossen werden.



## Summary

# Digital site identification and inventory of gravi-tational mass movements along the national trunk road network

## 1 Starting position and objective

Gravitational mass movements are widespread dangers in Germany and can cause considerable damage to the existing transport infrastructure. In addition to slope and embankment landslides as well as rockfalls, flow processes such as debris and sludge streams also belong to the category of mainly climatically controlled natural hazards. Close to traffic routes, they can not only cause structural damage, but also lead to serious traffic disturbance and thus to further economic costs.

Therefore, the aim of this research project is the development of a method that detects and classifies past gravitational mass movements, whereas the identification and analysis of the mapped events can serve as a basis for further approaches to monitor endangered slopes as well as to the early detection of potential cases of damage. The methodology has to be developed based on high resolution digital elevation models with 1 m or rather 2 m cell size. Furthermore, the client should be able to independently continue the work and add new districts as well as update existing areas once. Thus, all tools created for this purpose should be able to run in an ArcGIS desktop environment.

## 2 Literature analysis

The analysis of literature of the past two decades resulted in an accumulation of publications in the magazines Computers & Geosciences, Geomatics, Geomorphology, Natural Hazards as well as Risk and Natural Hazards. National, on the other hand, the search result is essentially limited to scientific qualification papers, technical reports or conference papers. The inspected publications were initially roughly divided into the three categories of landslides, rockfalls and flow processes at which a clear focus on landslide processes was observed. Publications regarding rockfall processes are mainly of geotechnical nature and are represented in a distinctly smaller quantity. Studies regarding flow or soil flow processes in terms of gravitational mass movement detection are almost not existent. The vast majority of the sighted publications deals with the creation of georisk and geohazard maps with the objective to detect potential landslide sites. Only in the last five to six years, the number of studies dealing with object-oriented methods for the identification of post-event gravitational mass movements on the basis of self-learning algorithms apparently increased. Regarding the validation results in these surveyed studies, rarely more than 70 % of the affected areas could be correctly classified as an event. Therefore, the visual validation still continues to play a key role in assessing the model quality.

## 3 Data acquisition and data processing

The following source data were available within the frame of the project:

- federal road network from the Federal Highway Information System, © BISStra/BASSt (2017),
- geological overview map GÜK200, © GeoBasis-DE/BKG (2013),
- digital landscape model ATKIS Basis-DLM, © GeoBasis-DE/BKG (2018),
- high resolution digital elevation models of the states North Rhine-Westphalia (1 m, © Land NRW (2018)), Saxony (2 m, © Land Sachsen (2018)) and Thuringia (1 m, © Land Thüringen (2018)),
- data from observed geohazards from different federal institutions.

The event data were first classified into the main process types of flow processes, rockfall processes and landslide processes and checked for plausibility with respect to their content or rather their attributed information. Furthermore, the data points were verified regarding their mapped position by the use of the high resolution digital elevation models. At this, approximately 68 % of flow processes and ca. 58 % of rockfall processes were classified as correct. Nevertheless, because the individual events could not be clearly depicted by the relief, mapping and recognition of individual rockfalls, rock erosion or similar processes

seems not to be feasible on the basis of the existing data set. Regarding the landslide processes, only ca. 15 % of the event data was confirmed by the comparison with the digital elevation models. This is a relatively low percentage and may be explained by its high dependency on size, age, soil or rock erosion resistance as well as the anthropogenic influence (e.g. clearing of landslide sites).

## **4 Concept**

### **4.1 Machine learning and software**

According to the results of the literature analysis, the development of the methodology will be based on supervised machine learning algorithms, using an object-oriented approach. For the detection of different gravitational mass movement events, the methods of artificial neural networks as well as random forests are considered. In regard to the technical solution, the applications Esri ArcGIS Desktop 10.5.1 (32-bit), Beak advangeo® (32-bit) and the Python 3 package Scikit-learn (64-bit) are available. The necessary data layers which are needed as the input for the machine learning methods are provided in form of morphological parameters, directly derived through calculations based on the high resolution digital elevation models.

### **4.2 Relief parameters**

The input data or rather parameters used in the sighted studies usually regard to models that were created for the purpose of georisk analysis - but not for the detection of structures. This is why primarily land use data and hydrological parameters as well as information on the soil or the existing geology were considered in these models. The relevant parameters for describing the surface conditions and structures in such models are often limited to roughness and position indices, slope angle, slope direction and also plan and profile curvatures.

Since these parameters alone do not provide enough scope to differentiate the surface and segment properties, further relief derivations from various sources will be used in this project with the target to describe the surface morphology as accurate as possible. To sum up, a total of 25 relief parameters based on calculations of the digital elevation models are available.

### **4.3 Pixel based vs. object-based approach**

Each pixel within a raster data set has a specific value (e.g. colour value, magnitude of a curvature parameter, slope, height value of the digital elevation model etc.), which is initially independent of the pixel values of its surrounding points. Thus, the pixels or rather their values do not have knowledge about the neighbourhood. However, particularly in gravitational mass movement events, the pixel values become more dependent to each other due to their collective rearrangement of material within a specific event (e.g. landslides). Hence, the common pixel-based approach, which is based on individual and independent pixels, is not suitable to reach the goals of this project. For this reason, an object-based approach seems to be the right choice. The creation of the required objects is realised in a first step by the selection of an appropriate relief parameter and second by the segmentation of this parameter using a geoinformation system (ArcGIS). This process aims to combine individual pixels to areas with similar characteristics as well as delineate areas with different surface structures from each other.

For the application in a machine learning approach, the objects generated this way are additionally assigned by statistical measures such as arithmetic mean, standard deviation and value range of any calculated relief parameter. Thus, the amount of available input data sets for the artificial neural network and random forest algorithms triplicates from 25 to 75 layers.

## **5 Training and model building**

### **5.1 Approach**

The result of the identification of gravitational mass movements based on their morphological properties depends on a large number of factors:

- the conditioning of the elevation model: filtering and smoothing,
- the used resolution of the input data,
- the included morphological surface parameters,
- the result of the segmentation process and object formation,
- the number and quality of mapped mass movements (study areas) as well as
- the selected machine learning algorithms.

Corresponding to this list of factors, there is a wealth of possible combinations for the initial model-building-process. To estimate the resulting impact of these factors, a large number of different test models was calculated and compared. The evaluation of the results was done by both visual verification of the revealed pattern as well as the comparison of correctly and incorrectly predicted areas.

At this point, one important finding is that compared to the 2 m digital elevation model, the use of the 1 m digital elevation model does not significantly improve the observed results. In return, the coarsening to 2 m cell size leads to a considerable decrease of calculation times as well as a distinctive reduction of the required storage space but results in a similar model performance. However, the available training sites seem to be the most important factor regarding the potential for improving the gained training results. Due to the small amount of existing training areas, it can be assumed that the models will be strongly improved by expanding the data set with more representative training sites. Furthermore, it could be shown that both the choice and the quantity of input relief parameters are also important factors with regard to the model quality. Test models with a very small number of input relief parameters (e.g. the ArcGIS function works only with 2 or less), did not reveal satisfying results.

## 5.2 Model building and training for state-wide application

### 5.2.1 Preliminary remarks

Due to the enormous data volume of the digital elevation models and their different derivatives, a creation of artificial neural network or even random forest models considering the whole state areas of North Rhine-Westphalia and Saxony was not feasible. Thus, a special training mosaic was built up by buffering each training site in each direction by 500 m and cutting out these areas. Finally, the snippets were assembled into one smaller mosaic. According to the number of training sites, the resulting mosaic contains 16 tiles (4 x 4) and has an extension of ca. 2.000 x 2.000 pixels (2 m cell size). Based on this high pixel-count, the models are limited to about 11 input layers due to technical limitations of the 32-bit software architecture. However, since the stock of usable training areas is not very extensive, the functionality of all processes and tools must ensure a much higher amount of training sites for the prospective application of the methodology. For this reason, identical models with a cell size of 5 m were additionally calculated. The coarsening from 2 m to 5 m raises the possible usable number of tiles to approximately 150, the conversion was carried out with the prepared input data sets from the 2-m-models.

Subsequently, the final training models were created for state-wide application to North Rhine-Westphalia and Saxony. To determine the optimal model settings, a bunch of different training scenarios was created, calculated and evaluated. Since the weighting of the input relief parameters is reinitiated during each training session, a total of five identical models for each resolution was generated (artificial neural network algorithm only).

### 5.2.2 Results

The comparison of the training results between the versions with 2 m and 5 m cell size, respectively, revealed no fundamental differences. This changes when comparing the results of the artificial neural network and random forest algorithms, in which partly distinctive differences were observed. In particular, this is expressed through the variation of the relative importance of single morphological input parameters, which not even vary between the used algorithms but even visibly alter amongst the artificial neural network models. Based on this, both the visual appearance as well as the ratio of correctly and incorrectly classified sites differ, resulting in apparent advantage for the models calculated by the random forest algorithm. Reasons for this may be:

- a) There are not enough training sites available.
- b) The summary of different mass movement processes to achieve a higher number of training areas is not conducive.

- c) The morphology is insufficiently described by the chosen relief parameters.
- d) The random forest algorithm may be the more appropriate method for classification.

Nevertheless, the artificial neural network algorithm achieves between 60 % and 70 % correctly classified training sites. For the random forests, the recognition rate of the training for the correctly identified areas reaches almost 100%. However, this may also be interpreted as an indication of so called model-overfitting.

## 6 Application

### 6.1 Preliminary remarks

For the application of the artificial neural network algorithm, the best two training models with the least error for both the 2 m as well as the 5 m versions, respectively, were selected and applied to the entire state areas. The results with a value range from 0 (event unlikely) and 1 (event probable) from both runs were averaged and then subsequently classified into the groups event/no event by an empirical determination of a threshold value close to 1 (e.g. 0.9). Since the used Python-3 package Scikit-learn offers a direct classification, only one model was calculated for each resolution version, making the application much more effective.

### 6.2 Results

The state-wide application of the 2 m versions with the artificial neural network algorithm lasts about 140 h for Saxony and about 410 h for North Rhine-Westphalia, which is clearly out of range in respect to a practicable application. In contrast, the random forest algorithm requires only 20 h for Saxony and 33 h for North Rhine-Westphalia, which is comparatively much less time. For the lower 5 m resolution versions, the computation time of the artificial neural network is dramatically decreased down to approximately 18 h for North Rhine-Westphalia and hence moved much closer to the random forest algorithm (ca. 12 h). At this resolution, both algorithms achieve acceptable calculation durations.

Furthermore, it was found that individual pixels, small groups of pixels and even partly anthropogenic structures such as narrow streets or roads, are classified as erosion events, so that the classification results have to be subjected by an additional correction.

Depending on the machine learning algorithm as well as the chosen resolution, significant differences in both the number and the position of the identified areas were observed. Therefore, the consideration of individual, single areas is not recommended, especially for relatively small sites with less than half a hectare. However, the results allow a comparison on a much smaller scale, as the models fundamentally detect events at least in the same regions. An improvement of the results should be achieved with additional event data.

### 6.3 Validation

The attribution and evaluation of the land use types showed that about 15 % to 20 % of the classified events are located in more or less anthropogenically influenced areas, which results in a significantly increased risk for misclassifications. In contrast, between about 65 % and 80 % of the areas are located primarily in forest areas, in which a lower misclassification rate can be assumed.

The comparison of the classification results with the average slope gradient showed a relatively high number of areas in classes with higher slope values, especially in the federal state of Saxony. This effect may be explained by the fact that in Saxony, for example in the Elbe Sandstone Mountains, many steep slopes are located in bedrock areas, which were identified by the models as landslide events due to their pronounced morphology. Although this represents basically a case of misclassification, it might indicate areas with active rockfall processes by inclusion of additional geological data, even if no single or rather individual rockfall event can be detected.

## 7 Conclusions and outlook

Although the creation, updating and improvement of the high resolution digital elevation models is continuously proceeded by the responsible authorities, update intervals of several years are not uncommon. Due to this reason, the development of a methodology on the basis of multi-temporal data was omitted. However, this approach would in principle be very well suited to map changes in digital elevation models, but for its realisation, only data with correspondingly short update intervals such as high-resolution satellite imagery can be used. Therefore, it is advisable to carry out a selective analysis of studies on this topic and, in principle, also for the diversity as well as the availability of these data and thereupon for the development of an alternative methodology to observe known individual events or rather monitoring purposes in general.

With regard to the results obtained in this research project, the availability of suitable training sites proved to be inadequate and hence can be seen as a limiting factor. Relating to the technical implementation of the algorithms, the random forest method seems to have a little advantage over the artificial neural network due to its shorter calculation duration as well as the possibility to work with the 2 m resolution data in a reasonable time frame. The use of the highest resolution available (1 m) does not seem very practical because the cost-benefit-ratio increases disproportionately in relation to the 2 m and 5 m cell size versions. Since the application results do not only differ between the algorithms but (in contrast to the training results) also between the 2 m and 5 m resolution versions, a valid assessment of individual events is only feasible by manual evaluation of pattern in the digital elevation models or alternatively by fieldwork.

Furthermore, it was shown that both the models of the artificial neural network algorithm and the random forest method generally identify the same regions, although the extent, the position and the number of individual sites are in detail not directly comparable. Moreover, potential areas with active rockfall processes might be indicated by the inclusion of additional geological data, particularly shown in the federal state of Saxony, for example in the Elbe Sandstone Mountains.

In conclusion, it can be stated that the basic procedure and methodology for the initial detection of mass movements on the basis of high-resolution digital elevation models was proven to be applicable, despite some misclassifications such as in anthropogenic areas. Nevertheless, more accurate results on the position and the general model quality do not seem achievable with the current data base, particularly with the least number of training sites. An application of the algorithm for monitoring purposes can almost be excluded due to the large update intervals and low availabilities of high-resolution digital elevation models.