

# InStRikE

Induced Seismicity Risk Estimation

<b>Programm / Ausschreibung</b>	Energieforschung (e!MISSION), Energieforschung, Energieforschung 7. Ausschreibung	<b>Status</b>	abgeschlossen
<b>Projektstart</b>	15.09.2021	<b>Projektende</b>	14.06.2023
<b>Zeitraum</b>	2021 - 2023	<b>Projektlaufzeit</b>	22 Monate
<b>Keywords</b>	Planungstool; Geothermie; Risikoabschätzung; induzierte Seismizität		

## Projektbeschreibung

Tiefe Geothermie erschliesst Reservoirs im Untergrund in Tiefen größer als 300 m. Interventionen in diesen Eroschichten sind auch mit Risiken verbunden. Da diese Art von erneuerbarer Energie vor allem in der Nähe von dicht besiedelten Städten effizient eingesetzt werden kann, ist eine Abschätzung des Risikos induzierter Seismizität in den Vordergrund gerückt. Das Projekt InStRikE setzt sich als Ziel, ein intelligentes Planungstool zu entwickeln, mit dem diese Risikoabschätzung bereits in einer sehr frühen Projektphase durchgeführt werden kann. Zunächst wird eine möglichst vollständige Datenbank von Risikoparametern von vorhandenen geothermischen Anlagen angelegt. Hierbei werden auch Messungen aus öffentlich zugänglichen Bohrlochdatenbanken herangezogen, um relevante Parameter abzuleiten. Im nächsten Schritt wird ein Modell entwickelt, in welchem diese Risikoparameter mit definierten geomechanischen Risikomechanismen in Zusammenhang gebracht werden. Durch die Anwendung von fortgeschrittenen Auswertungsmethoden, wie z.B. neuronalen Netzwerken, erhoffen wir uns auch neue Erkenntnisse um die Zusammenhänge zwischen geothermischen Operationen, geologischen und geomechanischen Gegebenheiten, und induzierter Seismizität besser verstehen und vorhersagen zu können.

## Abstract

Deep geothermal energy opens up underground reservoirs at depths greater than 300 m. Operations in these layers of the earth are also associated with risks. Since this type of renewable energy can be used efficiently, especially in the vicinity of densely populated cities, an assessment of the risk of induced seismicity becomes more important. The InStRikE project aims to develop an intelligent planning tool with which such risk assessment can be carried out at a very early stage of the project. First of all, a database of risk parameters as complete as possible is created from existing geothermal systems. Measurements from public available borehole databases will be used to derive relevant parameters. In the next step, a model is developed in which these risk parameters are related to defined geomechanical risk mechanisms. By using advanced evaluation methods, such as e.g. neural networks, we hope to gain new knowledge to better understand and predict the relationship between geothermal operations, geologic and geomechanical settings, and induced seismicity.

## Endberichtkurzfassung

InStRikE – Short summary

InStRikE aims to estimate the risk of induced seismicity at a geothermal site place, due to deep geothermal drilling and/or operation from the analysis of geological, geomechanical or operational parameters. These parameters were obtained from the literature, calculated from other parameters, or collected from adjacent or comparable sites. The goal is to use the prediction to optimize drilling and field development plans in a way to reduce the risk to induce or trigger seismicity, and to foster the utilization of deep geothermal plants.

Starting from a database compiled by Buijze et al. (2019), we have collected more parameters from more geothermal sites. All collected data underwent a rigid QC and QA to reach a high data quality for further processing, resulting in 100 (out of initial 158) geothermal operations chosen for further analyses. The data used predominantly represents geothermal operations from Europe, Oceania, and the United States of America, as these are better documented generally compared with other regions of the world. However, this also causes a bias in the geological/geothermal settings that are covered in the database, as operations from active rifting zones or volcanic settings might be under-represented as data from those regions are scarce or of low confidence. Since the seismicity observed during geothermal operations was used as part of the error criterion, the lack of a high-resolution seismic monitoring network since the start of an operation proved to be an exclusion criterion for inclusion in the statistical evaluation for many operations.

The data was then analyzed with advanced statistical methods, namely the Multi-Parameter-Influence Analysis (MPI) and the Logistic Regression Analysis (LOGIT). The MPI models worked with up to 15 parameters at the same time and were trained on 85 operations. Ten operations were used for evaluation of the method. The remaining five operations were used as a test set to validate and rank all models. The method assumes that parameters increasing the seismicity risk get a positive weighted score and parameters decreasing that risk get a negative weighted score. For each operation, the scores are summed and based on the sum it is estimated if an operation is prone to see seismicity above magnitude 2 ( $M>2$ ) or not. The LOGIT analysis works with subsets of four risk parameters, each of those subsets needs at least 15 geothermal operations to provide meaningful results. The method calculates probability scores and compares them with measured results (= observed seismicity below or above  $M2.0$  in this case). If the resulting likelihood is above or below a certain threshold the model's prediction is either right or wrong.

Both chosen methods can successfully estimate the seismicity observed during the operations. For three risk parameter sets (each containing 15 parameters) and a blind test set of count 5 the MPI method predicts between 84% and 100% successfully. For the LOGIT analysis, four independent trainings were evaluated using three different criteria are set: setting the threshold of probability to 0.5, 89% of the operations are correctly predicted either as seismic ( $> M2.0$ ) or non-seismic ( $\leq M2.0$ ). Applying two thresholds, either smaller than 0.4 probability for non-seismic and higher than 0.6 for seismic operations, 94% of the operations are predicted correctly. The third criterion, using slightly different thresholds, lower than 0.3 and greater than 0.7 probability, results in a 93% correct prediction.

Given the small overall dataset ( $n=100$ ), results of both analyses are considered good and encouraging to follow the applied methodologies further on. We are confident that the existing results are suitable to raise interest with geothermal operators to share their data with us for further improvement of the methodology and/or to apply the InStRikE method for their planned geothermal projects.

Ziel von InStRikE ist es, das Risiko induzierter Seismizität bei Geothermieanlagen aufgrund von Tiefengeothermiebohrungen und/oder des Anlagenbetriebes anhand der Analyse geologischer, geomechanischer und betrieblicher Parameter abzuschätzen. Diese Parameter wurden aus der Literatur entnommen, aus anderen Parametern berechnet oder von benachbarten oder vergleichbaren Standorten gesammelt. Ziel ist es, die Abschätzung zur Optimierung von Bohr- und Reservoirentwicklungsplänen zu nutzen, um das Risiko von induzierter Seismizität zu verringern und die Nutzung von Tiefengeothermieanlagen zu fördern.

Ausgehend von einer von Buijze et al. (2019) publizierten Datenbank über Geothermieanlagen weltweit haben wir diese um zusätzliche Geothermiestandorte und Parameter erweitert. Alle gesammelten Daten wurden einer strengen Qualitätskontrolle und Qualitätssicherung unterzogen, um eine hohe Datenqualität für die weitere Verarbeitung zu gewährleisten. Dies führte dazu, dass 100 (von ursprünglich 158) geothermische Betriebe für die weitere Analysen ausgewählt wurden. Die verwendeten Daten repräsentieren überwiegend Geothermieanlagen aus Europa, Ozeanien und den Vereinigten Staaten von Amerika, da diese im Vergleich zu anderen Regionen der Welt allgemein besser dokumentiert sind. Allerdings führt dies auch zu einer Verzerrung der in der Datenbank abgedeckten geologischen/geothermischen Umgebungen, da Operationen aus aktiven Rifting-Zonen oder vulkanischen Umgebungen möglicherweise unterrepräsentiert sind, da die Daten dieser Regionen nur knapp oder von geringer Zuverlässigkeit sind. Da die beobachtete Seismizität während des Anlagenbetriebs im Fehlerkriterium verwendet wurde erwies sich für viele Standorte das Fehlen eines hochauflösenden seismischen Überwachungsnetzes seit Beginn des Betriebs als Ausschlusskriterium für die Aufnahme zur statistischen Auswertung.

Anschließend wurden die Daten mit fortschrittlichen statistischen Methoden analysiert, nämlich der Multi-Parameter-Influence-Analyse (MPI) und der Logistischen Regressionsanalyse (LOGIT). MPI-Modelle arbeiten mit bis zu 15 Parametern gleichzeitig und wurden mit 85 Datenbankeinträgen (geothermischen Projekten) trainiert. Zur Evaluierung der Methode wurden zehn Projekte verwendet. Die verbleibenden fünf Projekte wurden als Testsatz zur Validierung und zum Gewichten aller trainierten Modelle verwendet. Die Methode basiert darauf, dass ein Parameter, der das Seismizitätsrisiko erhöht, einen positiv gewichteten Wert erhält und ein Parameter, der dieses Risiko verringert, erhält einen negativen gewichteten Wert. Für jedes Projekt wurden die Ergebnisse summiert und gewichtet und anhand der Summe abgeschätzt ob die Projekte zur Seismizität grösser Magnitude 2 ( $M > 2$ ) neigt oder nicht. Die LOGIT-Analyse arbeitet mit Teilmengen von vier Risikoparametern, wobei jede davon mindestens 15 geothermische Projekte benötigt, um aussagekräftige Ergebnisse zu liefern. Die Methode berechnet Wahrscheinlichkeitswerte und vergleicht diese mit gemessenen Ergebnissen (= beobachtete Seismizität unter oder über  $M 2$  in diesem Fall). Wenn die resultierende Wahrscheinlichkeit über oder unter einem bestimmten Schwellenwert liegt, ist die Vorhersage des Modells entweder richtig oder falsch.

Mit beiden gewählten Methoden kann die während des Betriebs beobachtete Seismizität erfolgreich abgeschätzt werden. Für drei Risikoparametersätze (je 15 Parameter) und einem Blindtestsatz von 5 Projekten prognostiziert die MPI-Methode zwischen 84 % und 100 % erfolgreich. Für die LOGIT-Analyse wurden vier unabhängige Trainings anhand dreier verschiedener Kriterien bewertet: Durch die Festlegung der Wahrscheinlichkeitsschwelle auf 0,5 werden 89 % der Vorgänge korrekt vorhergesagt, entweder als seismisch ( $> M 2,0$ ) oder nicht seismisch ( $\leq M 2,0$ ). Durch die Anwendung von zwei Schwellenwerten, die entweder kleiner als 0,4 Wahrscheinlichkeit für nichtseismische Vorgänge und höher als 0,6 für

seismische Vorgänge sind, werden 94 % der Vorgänge korrekt vorhergesagt. Das dritte Kriterium, das leicht unterschiedliche Schwellenwerte mit einer Wahrscheinlichkeit von weniger als 0,3 und mehr als 0,7 verwendet, führt zu einer zu 93 % korrekten Vorhersage.

Angesichts des geringen Gesamtdatensatzes (n=100) werden die Ergebnisse beider Analysen als gut und ermutigend angesehen, die angewandten Methoden weiter zu verfolgen. Wir sind zuversichtlich, dass die vorliegenden Ergebnisse geeignet sind, das Interesse von Geothermiebetreibern zu wecken, ihre Daten mit uns zu teilen, um die Methodik weiter zu verbessern und/oder die InStRikE-Methode für ihre geplanten Geothermieprojekte anzuwenden.

## **Projektkoordinator**

- DI Julia Dießl

## **Projektpartner**

- GeoMechanics Technologies
- NiMBUC Geoscience OG