© 2026 Svenska Bergteknikföreningen och författarna/Swedish Rock Engineering Association and authors.

HUR KAN VI UPPSKATTA MARKVIBRATIONER MED MASKININLÄRNING?

HOW TO ESTIMATE GROUND SURFACE VIBRATIONS USING MACHINE LEARNING?

Staffan Risberg, ITASCA
Mikael Svartsjaern, ITASCA
Jonny Sjöberg, ITASCA
Helena Zakrisson, LKAB

SAMMANFATTNING

Alla svenska gruvor, och även de flesta internationella, måste förhålla sig till strikta markvibrationskrav i form av fastställda miljövillkor. Traditionellt är dessa markvibrationer i huvudsak kopplade till sprängningsinducerade vibrationer, men för seismiskt aktiva gruvor står idag seismiken för en allt större andel av de höga vibrationerna. Att koppla samman seismik underjord med resulterande markvibrationer ovanjord utgör därmed ett viktigt steg mot att uppskatta den miljöpåverkan som framtida gruvbrytning medför.

Denna artikel beskriver hur gruvinducerad seismik i Kiirunavaaragruvan kopplas till uppmätta markvibrationer i Kiruna stad. Med hjälp av maskininlärningsmodeller associeras seismikdata med resultat från numerisk modellering för att generera uppskattade seismiska händelser, vilka sedan används för att prognosticera markvibrationer i Kiruna stad. I artikeln presenteras även resultat från verifiering av modellerna där prognosticerade markvibrationer från uppskattade seismiska händelser för perioden 2018–2024 jämförs med uppmätta värden för samma tidsspann.

Nyckelord: gruvseismik, markvibrationer, prognostisering, maskininlärning, numerisk modellering

SUMMARY

All mines must consider strict ground vibration limits defined by environmental permits. Traditionally, these vibrations have mainly been associated with blasting, but in seismically active mines, mining-induced seismicity now accounts for an increasingly larger share of the high vibration levels. Establishing a link between underground seismic events and resulting ground surface vibrations is therefore essential when assessing the environmental impact of future mining activities.

This article presents a method for relating mining-induced seismicity in the Kiirunavaara mine to recorded ground surface vibrations in the town of Kiruna. Using machine learning models, seismic data are combined with numerical modelling results to generate synthetic seismicity, which is subsequently used to forecast ground surface vibrations in Kiruna. The article also presents verification of the described method, where predicted ground surface vibrations from synthetic seismic activity for the period 2018–2024 are compared with actual recorded values from the same timeframe.

Keywords: mine seismicity, ground surface vibrations, forecast, machine learning, numerical modeling

1 INTRODUKTION

För att övervaka uppkomna markvibrationer som uppföljning av de krav som ställs på LKAB via miljödom har LKAB under många år haft ett omfattande vibrationsövervakningssystem både under jord och på markytan. Som ett resultat av detta finns idag stora databaser innehållande seismik- och markvibrationsdata, vilka kan användas som träningsdata för maskininlärningsmodeller. Syftet med denna artikel är att demonstrera en metod för att prognosticera markvibrationer från framtida gruvbrytning, specifikt från inducerad seismik, med hjälp av maskininlärningsmodeller. Maskininlärning utgör ett utmärkt verktyg för att prognosticera markvibrationer från seismik från ett statistiskt perspektiv utan att explicit prognosticera enskilda seismiska händelser.

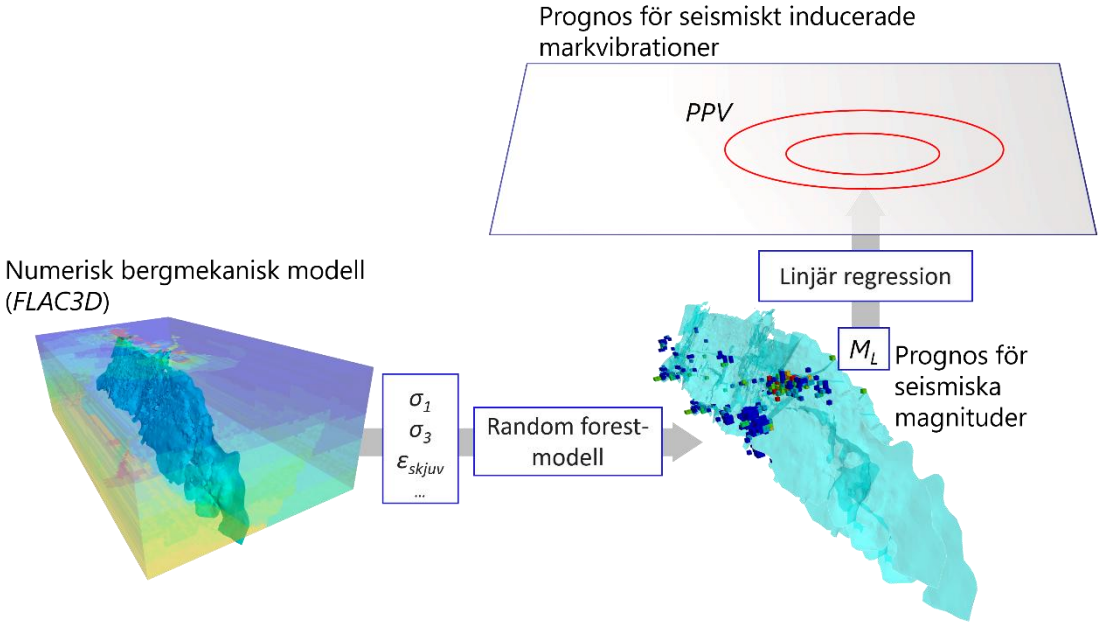
2 METOD OCH DATA

2.1 Övergripande metod

I Figur 1 visas övergripligt den metod som använts för att prognosticera markvibrationer med utgångspunkt från en numerisk deformationsmodell. För denna studie har två maskininlärningsmodeller nyttjats: en Random forest-modell och en linjär regressionsmodell. Det framtagna arbetsflödet kan beskrivas enligt följande:

1. Simulering av rasbrytning utförs i en FLAC3D-modell till dagens datum.
2. Beräknade spänningar, töjningar och energier plockas ut ur FLAC3D-modellen och matas in i en Random forest-modell som identifierar samband mellan numeriska utdata och uppmätt seismik under motsvarande tidsperiod.
3. Random forest-modellen körs sedan med data från den numeriska modellen för att skapa syntetisk seismik, som sedan jämförs med faktiskt uppmätt seismisk aktivitet för att verifiera modellens randvillkor.
4. Prognosticerad seismik används därefter som indata i en linjär regressionsmodell som räknar ut motsvarande markvibrationer baserat på seismisk lokal magnitud M_L och avstånd från källan. Prognosticerade markvibrationer jämförs sedan mot faktiskt uppmätta vibrationer i syfte att validera arbetsmetodiken inför prognos för framtida brytning.

Enligt ovan arbetssätt krävs att den linjära regressionsmodellen och Random forest-modellen tränas, det vill säga lär sig samband utifrån historiska data för att sedan extrapolera dessa samband in i framtiden.



Figur 1 Flöde som översiktligt beskriver använd metodik. σ_1 (största huvudspänning), σ_3 (minsta huvudspänning) och ε_{skjuv} (skjuvtöjning) är ett urval av indataparametrar för statistisk prognos av seismiska händelser (M_L).

2.2 Maskininlärning

Det finns ett flertal olika maskininlärningsalgoritmer (exempelvis linjär regression, neurala nätverk och k-nearest neighbors) som lämpar sig olika bra beroende på data och problemställning. Dessa kan delas in i tre huvudkategorier:

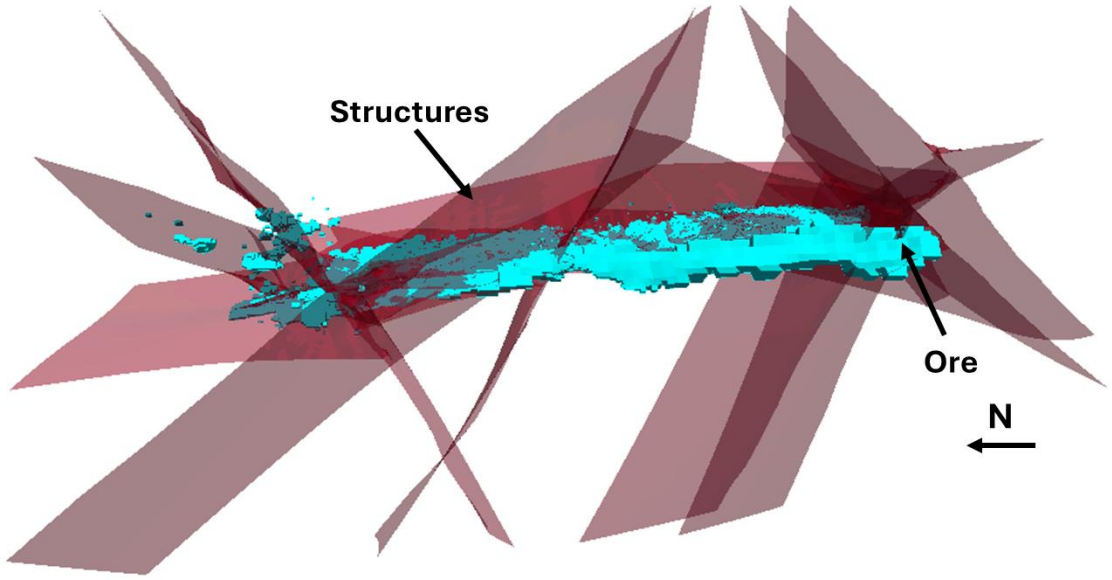
- övervakad inlärning,
- oövervakad inlärning, och
- förstärkt inlärning.

Skillnaden mellan metoderna ligger i hur data är strukturerad. I många fall är data märkt, vilket innebär att det är tydligt vad som är en observation (exempelvis en seismisk magnitud) och vad som är utfallet (exempelvis resulterande markvibrationer). I dessa fall är övervakad inlärning ett naturligt val, som exempelvis en enkel linjär regression. Oövervakad inlärning används då data inte är märkta, till exempel för att identifiera kluster, och förstärkt inlärning avser modeller som lär sig genom att prova handlingar och justera sitt beteende baserat på återkoppling i form av belöning eller bestraffning (Domo, 2025). I metoden som beskrivs i denna artikel tillämpas övervakad inlärning, där både linjär regression och Random forest-modeller tillhör denna kategori.

2.3 *Numeriska indata och träningsdata*

Indata till den numeriska modellen bestod av en geologisk modell med geomekaniska enheter och storskaliga strukturer där storskaliga strukturer (se Figur 2) simulerades implicit i ett kontinuum. FLAC3D-modellen beskrivs mer ingående av Svartsjaern m.fl. (2025) och beskrivs därför inte vidare i denna artikel.

Som träningsdata till maskininlärningsmodellerna användes en markvibrationsdatabas och gruvans seismiskdatabas. Markvibrationsdatabasen innehöll uppmätta vibrationer uttryckt i *PPV* (Peak Particle Velocities) för ett flertal marksensorer i Kiruna stad med koordinater och tidpunkt för respektive mätning. Seismikdatabasen innehöll data för ett stort antal seismiska händelser uppmätta mellan 2010 och 2024 med varierande seismisk magnitud (M_L) och koordinater för respektive händelse.



Figur 2 Vy ovanifrån visande modellerade storskaliga strukturer (röda) i FLAC3D i relation till malmkoppen (blå).

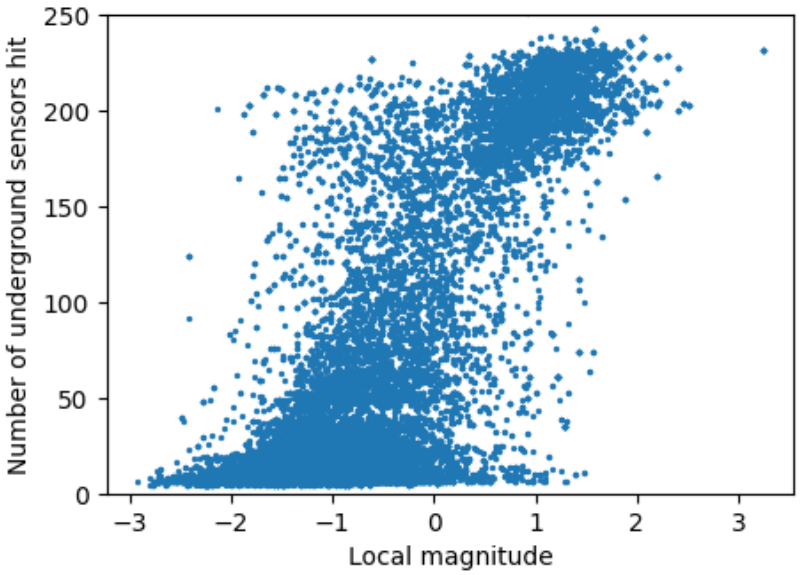
2.4 *Bearbetning av träningsdata*

Noggrann bearbetning av träningsdata är viktigt för att undvika att maskininlärningsmodeller lär sig felaktiga mönster. Seismikdatabasen och markvibrationsdatabasen var initialt helt separerade vilket innebar att specifika seismiska händelser inte var kopplade till specifika markvibrationer. Att enbart tidsfiltrera varje seismisk händelse och koppla till samtliga uppmätta markvibrationer inom ett visst tidsfönster visades medföra en stor mängd felaktiga sammankopplingar, alltså att uppmätta markvibrationer kopplats till osannolika seismiska källor. I detta fall antogs felen orsakas av att (i) seismiska händel-

ser kopplats till markvibrationer orsakade av icke-seismiska källor (exempelvis trafik) samt att (ii) flera olika seismiska händelser kunde kopplats till en och samma markvibration. Detta eftersom tidstämplarna i under- och ovanjordssensorerna inte varit synkroniserade. För att minska antalet felkopplingar applicerades ett antal filter vid koppling mellan de två databaserna:

- All data registrerat innan 2013 uteslöts ur analysen. Detta på grund av att en större utbyggnad av det seismiska övervakningssystemet genomfördes 2009–2012 (Dahnér m.fl., 2012) vilket tydligt påverkade systemets upplösning.
- Tidsfönstret mellan en seismisk händelse och motsvarande markvibration bör inte överskrida ett orimligt intervall. Genom att testa olika tidsfönster utvärderades hur detta påverkade träffsäkerheten. Ett tidsfönster på 0–2 s mellan respektive seismiska händelse och motsvarande uppmätta markvibrationer gav bäst noggrannhet och kan även anses rimligt med tanke på de fysiska avstånden efter att synkroniseringsfelen beaktats.
- Enbart större seismiska händelser bör resultera i signifikanta utslag på markvibrationssensorer. Av denna anledning inkluderades endast de händelser som registrerats av ett relativt större antal sensorer under jord eftersom dessa händelser med stor säkerhet bör ha gett utslag på flertalet marksensorer. Detta motiveras av de två distinkta klustren i Figur 3.
- Markvibrationer från seismiska händelser registreras sannolikt av flertalet marksensorer samtidigt (inom 0–2 s). Genom att anta att varje seismisk händelse måste registreras av minst 5 marksensorer inom 0–2 s antogs större delen av icke-seismiska markvibrationer kunna filtreras bort.

Genom att nyttja ovan filter vid kopplingen mellan den seismiska databasen och markvibrationsdatabasen kunde en trovärdig kombinerad databas skapas och användas för träning av maskininlärningsmodellerna. Databasen omfattade tidpunkt och uppmätt svängningshastighet PPV från respektive marksensor, tillsammans med tidpunkt och seismisk magnitud M_L för respektive seismiska händelse som antogs orsaka responsen på markytan, se Tabell 1.



Figur 3 Data för seismiska händelser med uppmätt markvibration inom 0–2s. För varje händelse plottas seismisk magnitud och antal underjordssensorer som registrerat respektive händelse. Endast händelser som registrerats av minst 150 underjords-sensorer beaktades.

Tabell 1 Utdrag ur den korrelerade databasen innehållande uppmätt seismisk magnitud M_L och motsvarande uppmätt markvibration PPV för ett antal seismiska händelser.

Datum (markyta)	***PPV*** **[mm/s]**	**Datum (under jord)**	M_L	**Avstånd [m]**	**Tidsskillnad [s]**
2023-08-10 20:11:16	0.25	2023-08-10 20:11:14	1.01	2430.0	1.56
2023-08-20 01:14:34	0.15	2023-08-20 01:14:33	0.93	2436.0	0.14
2023-09-01 17:38:58	0.50	2023-09-01 17:38:57	1.19	2198.0	0.66
2023-09-08 16:15:56	0.20	2023-09-08 16:15:55	1.28	2593.0	0.19
2023-09-11 02:15:05	0.30	2023-09-11 02:15:03	1.20	2277.0	1.32

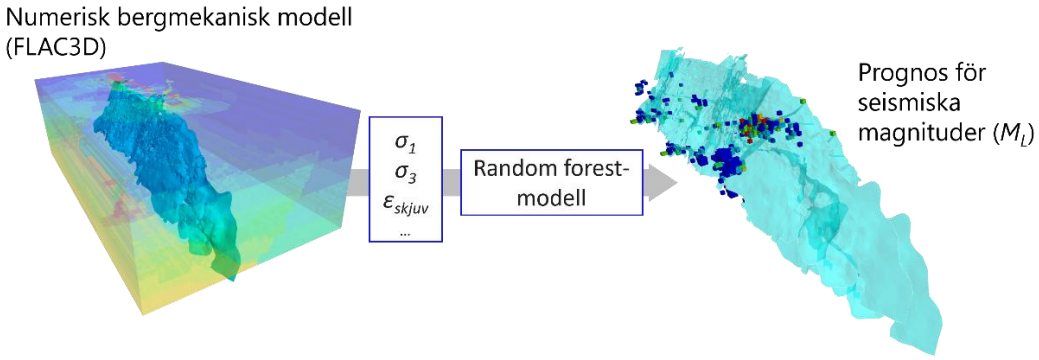
3 MASKININLÄRNINGSMODELLER

3.1 Maskininlärningsmodell för markvibrationsprognos

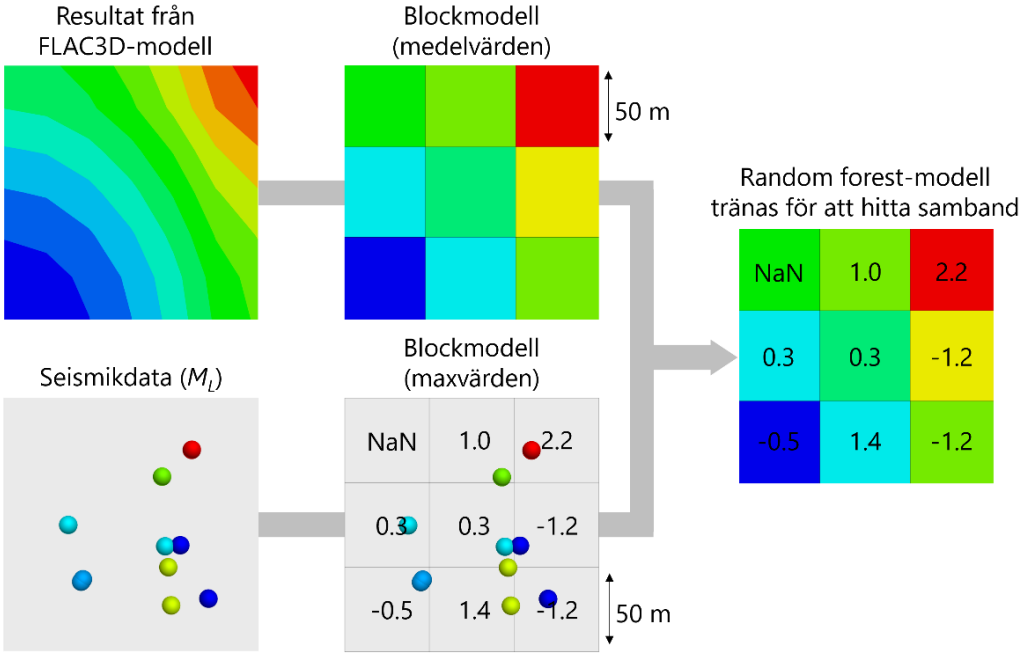
Den kopplade seismik-och-markvibrationsdatabasen (Tabell 1) användes för att träna en maskininlärningsmodell för att uppskatta *PPV* från en godtycklig seismisk händelse i form av ett magnitud-avståndssamband. Det förväntas att ett sådant samband bör vara relativt enkelt: ett kortare avstånd och högre seismisk magnitud bör ge ett högre *PPV*-utslag än ett större avstånd med lägre magnitud. Detta motiverade valet av en linjär regressionsmodell.

3.2 *Maskininlärningsmodell för prognos av seismisk magnitud*

Den numeriska FLAC3D-modellen beräknade inte seismiska magnituder explicit, i stället tränades en andra maskininlärningsmodell att hitta samband mellan observerade seismiska händelser och grundläggande FLAC3D-resultat (spänningsmagnituder, töjningar och energier), se Figur 4. Träningsdata för denna maskininlärningsmodell togs fram genom att simulera periodvis brytning fram till och med år 2024 i FLAC3D och kombinera modellresultat för respektive period med seismikdata för motsvarande period. De två datatyperna kombinerades i en blockmodell med 50x50x50 m upplösning enligt Figur 5.

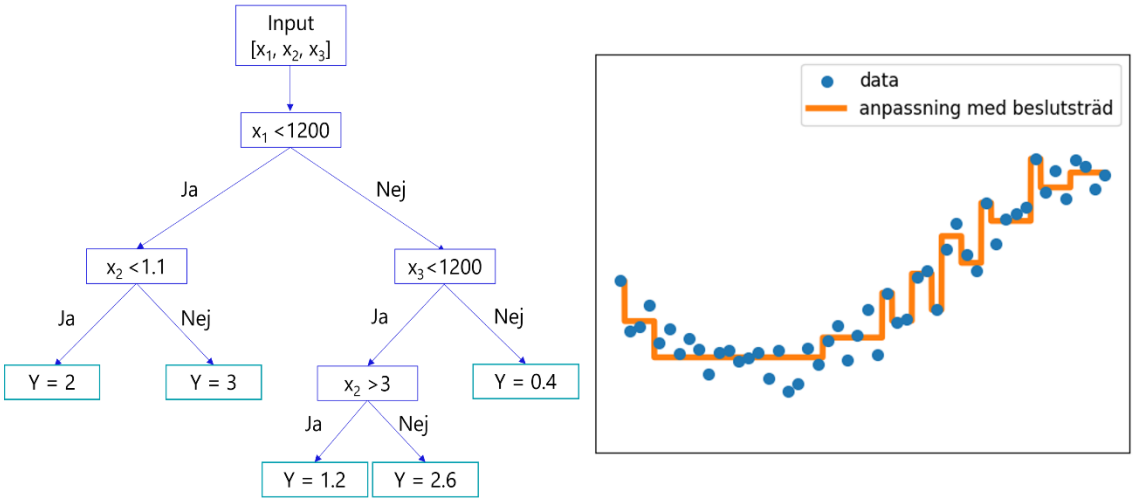


Figur 4 Visualisering av principen för prediktion av seismik från FLAC3D-resultat med hjälp av efterprocessering i en Random forest-modell.

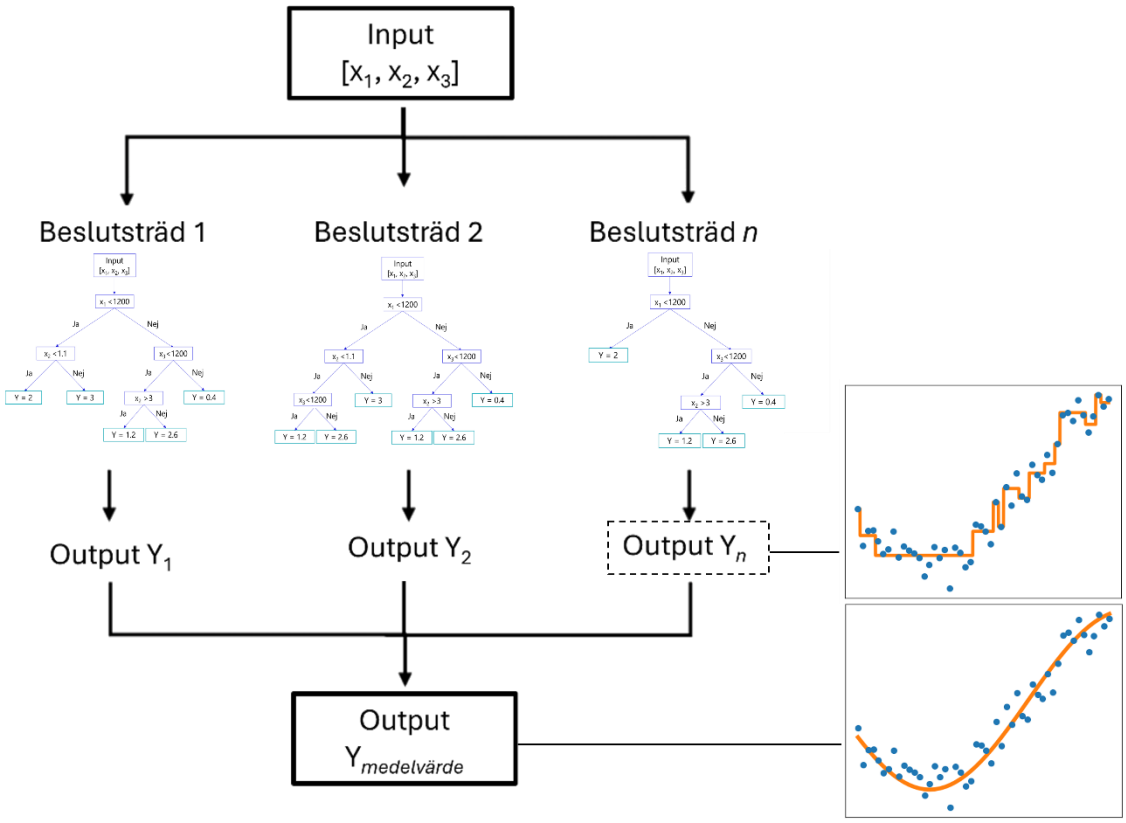


Figur 5 Sammanställning av FLAC3D-data och seismikdata i en gemensam blockmodell. Genom att jämföra numeriska modellresultat (spänningsmagnituder, töjningar och energier) mot seismikdata för motsvarande period (kvartal) kan dessa relateras till varandra i en maskininlärningsmodell.

Seismik är komplext och samband mellan indata och utdata (orsakssamband) är därför svåra att explicit definiera. En ökning eller minskning av viss indataparameter behöver inte nödvändigtvis ge motsvarande ökning eller minskning av den seismiska potentialen, utan kanske bara i vissa kombinationer med andra parametrar. En Random forest-modell ansågs lämplig för efterprocessering eftersom denna typ av modell bygger på beslutsträd vilka inte förutsätter direkta samband mellan indata och utdata. Modellen delar in data i kategorier beroende på hur indata och utdata verkar höra ihop (med succesiva ja/nej-satser) (Raschka & Mirjalili, 2019), se Figur 6. På så sätt är beslutsträd-modeller betydligt mer flexibla jämfört med till exempel linjära modeller. En nackdel med flexibiliteten är att modellerna lätt kan överanpassas, alltså att de anpassas efter brus vilket också illustreras i Figur 6. Genom att träna flera olika beslutsträd och ta medelvärdet från resultaten fås ett mer generellt resultat, vilket utnyttjas i Random forest-modeller (Figur 7).



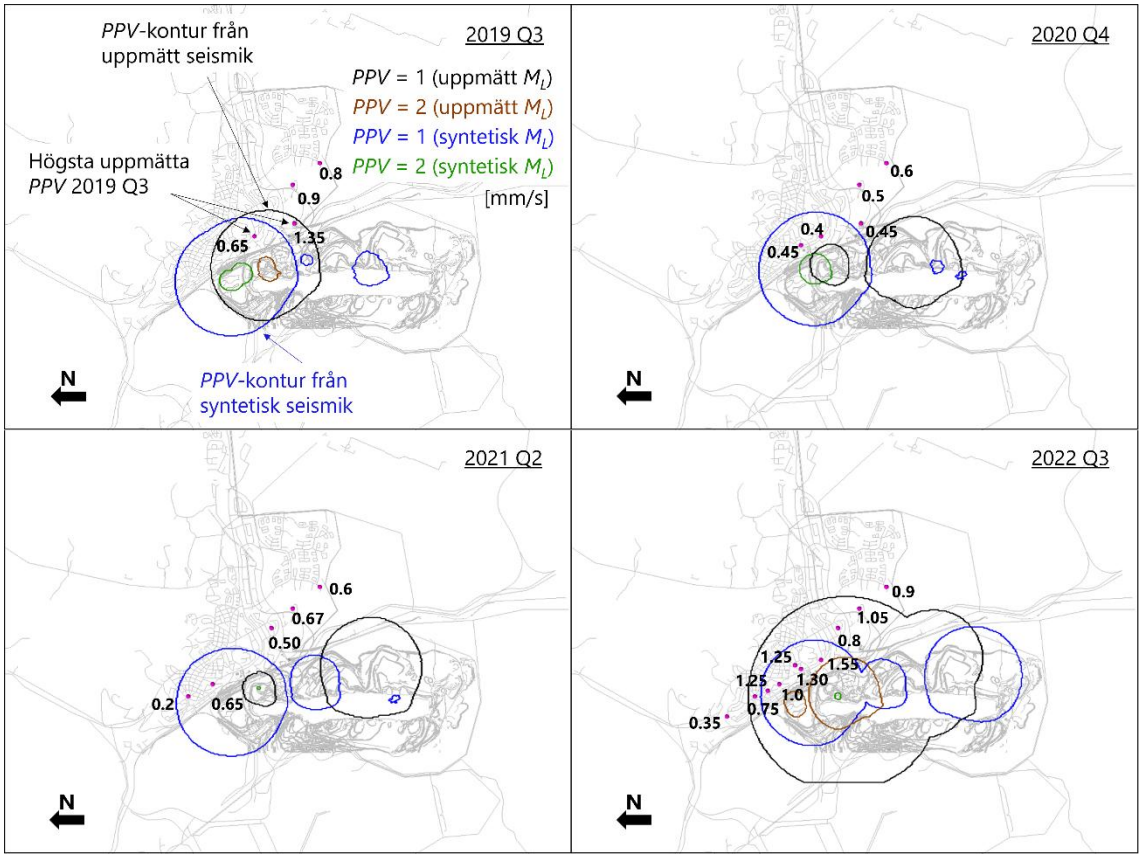
Figur 6 Till vänster visas ett beslutsträd som med hjälp av x_1, x_2, x_3 predikterar Y. I grafen till höger visas hur ett (annat) beslutsträd anpassats till data.



Figur 7 Random forest-modeller väger samman lösningen från flertalet olika beslutsträd (ofta i hundratals eller tusentals). Genom att ta medelvärdet av alla lösningar fås en mer robust lösning med lägre risk för slumpmässiga fel. Figur modifierad från Chu m.fl. (2024).

4 VALIDERING AV MODELLER

Modellerna validerades genom att skapa kvartalsvisa *PPV*-konturkartor för år 2014–2025 innehållande två olika konturer – en med beräknade markvibrationer från prognosticerad (syntetisk) seismik och en med beräknade markvibrationer från uppmätta M_L. Dessa två konturuppsättningar jämfördes mot faktiska *PPV*-mätningar från markvibrationsdatabasen för samma tidsperiod, se exempel i Figur 8.



Figur 8 Validering av maskininlärningsmodellerna för 4 kvartal mellan år 2019–2022. Konturer och värden motsvarar de högst uppmätta/beräknade PPV för respektive kvartal (i mm/s).

Generellt ansågs de två *PPV*-konturerna stämma relativt väl överens, både internt och mot uppmätta *PPV* > 1 mm/s. Något som framgår av figuren (Figur 8) är den stora variation i uppmätta *PPV*-värden där flertalet närliggande marksensorer visar, mellan varandra, mycket skilda värden, särskilt för Q3 2022. Detta indikerar kvarvarande brus i data som inte filtrerats bort i databearbetningen. Att modellerna inte anpassats efter detta brus kan anses positivt då det antyder att modellerna är generella och inte överanpassade på brus i träningsdata.

5 DISKUSSION OCH LÄRDOMAR

Det är viktigt att förstå att maskininlärningsmodeller enbart kan sammankoppla in- och utdata och det finns ingen garanti för att dessa kopplingar har någon fysikalisk mening (kausalitet). Detta är särskilt viktigt för Random forest-modeller där det är svårt att förutsäga vilka regler modellen kommer att applicera på data. Därför är det ett måste att för dessa analyser ha bra insikt i de data som modellen tränats på för att kunna förstå och kritiskt granska modellens prediktioner.

Det finns idag många olika maskininlärningsalgoritmer att välja mellan men det är komplext att på ett tillförlitligt och systematiskt sätt jämföra utfallet och motivera val av algoritm mellan exempelvis Random forest och neurala nätverk. För komplexa problem i praktiska tillämpningar som detta påverkas resultaten från jämförelser mellan modeller i hög grad av slumpmässiga variationer eller överanpassning av modeller (Kitchin, 2020). Valet av en Random forest-algoritm och linjär regression kan i detta fall motiveras av algoritmernas enkla uppbyggnad och av att de resultat som erhålls är relativt enkla att förklara, och därmed granska.

Sammanfattningsvis visar den demonstrerade metoden att kombinationen av maskininlärning och numerisk modellering möjliggör nyttjande av historiska data för att kunna prognosticera markvibrationer med numeriska modeller utan att explicit modellera seismiska händelser. Därmed är det inte nödvändigt att behöva definiera komplexa samband mellan dataparametrar vilka kan antas leda till seismik utan i stället låta maskininlärningsmodeller hitta och nyttja korrelationer i data.

TACK

Författarna vill rikta ett stort tack till LKAB för möjligheten att publicera innehållet i denna artikel och även till Jimmy Töyrä för hans hjälp med att definiera ramen för arbetet.

REFERENSER

Chu, T., Nguyen, T. C., Yoo, H., & Wang, J. 2024. ESP vibration prediction based on pump operating conditions in laboratory using machine learning. *Geoenergy Science and Engineering*, 240, 213039.

Dahnér, C., Malmgren, L. & Bošković, M. 2012. Transition from non-seismic mine to a seismically active mine: Kiirunavaara Mine. ***Rock Engineering and Technology for Sustainable Underground Construction (Proceedings, Eurock 2012 — ISRM International Symposium, Stockholm, May 28–30, 2012)***. Stockholm: BeFo. 2012.

Domo. 2025. *What are Machine Learning Basics?* https://www.domo.com/glossary/what-are-machine-learning-basics [Hämtad 2025-01-23]. Domo. USA.

Kitchin. 2020. *Nonlinear models in sklearn*. https://kitchingroup.cheme.cmu.edu/s20-06681/08-nonlinear-sklearn/08-nonlinear-sklearn.html [Hämtad 2025-01-26]. Carnegie Mellon University. USA.

Raschka, S., & Mirjalili, V. 2019. ***Python Machine Learning Third Edition***. Packt Publishing Ltd., ISBN 978-1-78995-575-0, Chapter "Decision tree regression".

Svartsjaern, M., Rentzelos, T., Risberg, S., Sjöberg, J., Zakrisson, H., Thielsen, C. and Furtney, J. 2025. Ground Surface Vibration Estimated from Seismic Potential Using Machine Learning, in ***Proceedings of the 11th International Symposium on Rockbursts and Seismicity in Mines (RaSim, 2025)***, pp 318–336. 24–29 August; Luleå, Sweden. 2025.