Mine Closure 2026 – AB Fourie, G Boggs, J Heyes & M Tibbett (eds)

doi:10.36487/ACG_repo/2615_80

The application of LiDAR and machine learning in managing historical mining features in Victoria, Australia

Shannon Herley [a,*], Charlotte Riley [a], Luong Tran [a], Louise Goldie Divko [a], Rami Eid [b], Elizabeth Silver [c], Long Dang [c]

[a] *Geological Survey of Victoria, Resources Victoria, Department of Energy, Environment and Climate Action, Australia*

[b] *Geoscience Australia, Australia*

[c] *WSP, Australia*

Abstract

Accurate mapping of historical mining features is critical for their management, providing essential information for a wide range of stakeholders – from local communities and land managers planning day-to-day activities to emergency services responding to incidents. This study presents a large-scale machine- learning workflow for detecting and classifying previously unrecorded legacy mining features using Light Detection and Ranging (LiDAR) data and digital elevation models (DEMs).

Two detection methods were developed to identify potential mining features: a sink detector which identifies hydrologic sinks in the DEM, and a noise detector which identifies clusters of noise-classified points in the LiDAR point cloud. Subsets of these detections were manually labelled by humans as 'mine shaft', 'surface workings' or 'not sure/NA/other' using visual inspection. These were used to train and test the machine learning image classification model. Statistical and spatial filtering techniques were then applied to refine the results.

The method was applied as part of a broader program to establish a centralised and comprehensive database of mining features using available historical records. Its implementation has increased the number of recorded legacy mining features in Victoria's historical database from approximately 36,000 to more than 250,000. These results demonstrate the value of using LiDAR data, DEM and machine learning to improve the identification and management of historical mining features, and provide a rapid, scalable approach for jurisdictions to better map and manage legacy mining features.

Keywords: *database, legacy, LiDAR, machine learning, mine shafts*

1 Introduction

The State of Victoria in southeast Australia has a mining history that extends back to the goldrush of the 1850s. The legacy of the early miners is concentrated in central Victoria and to the state's northeast, where over 80 million ounces of gold has been produced to date. As part of a broader study to establish a framework to manage abandoned and legacy mines and quarries on Crown land, existing databases were consolidated (Riley et al. 2025). Field observations of existing features in multiple locations highlighted that many legacy mining features in the landscape were not recorded or characterised.

Light Detection and Ranging (LiDAR) has been used to identify and map mining features (e.g. Gergelova et al. 2021), particularly where those features are not easily accessible or visible on satellite imagery. LiDAR is a remote sensing method that uses light from a pulsed laser on an aircraft to measure distances to the Earth (National Oceanic and Atmospheric Administration 2024). The pulses are used to calculate the distance between the laser scanner and the Earth's surface to generate a precise, 3D surface representation in the

[*] Corresponding author. Email address: shannon.herley@deeca.vic.gov.au

form of a point cloud (U.S. Geological Survey 2025). Points are classified into surface-type classes (e.g. buildings, vegetation, ground, water or noise). Ground-classified points are used to generate digital elevation models (DEMs), while noise-classified points represent erroneous outliers that do not correspond to real surface features.

The aim of this study was to use LiDAR data coupled with machine learning to automate a process to detect and classify previously unrecorded mining features.

2 Methods

2.1 Study areas

The study areas (Figure 1) were selected over regions with the most known recordings of legacy mining features in conjunction with the presence and resolution of LiDAR data.

Study area 1 (Golden Plains), study area 2 (Greater Melbourne and Central Goldfields), and study area 3 (Cobungra-Dargo) are introduced in 3 summary reports (Silver et al. 2025a, 2025b, 2026) that accompany the final datasets of each study area.

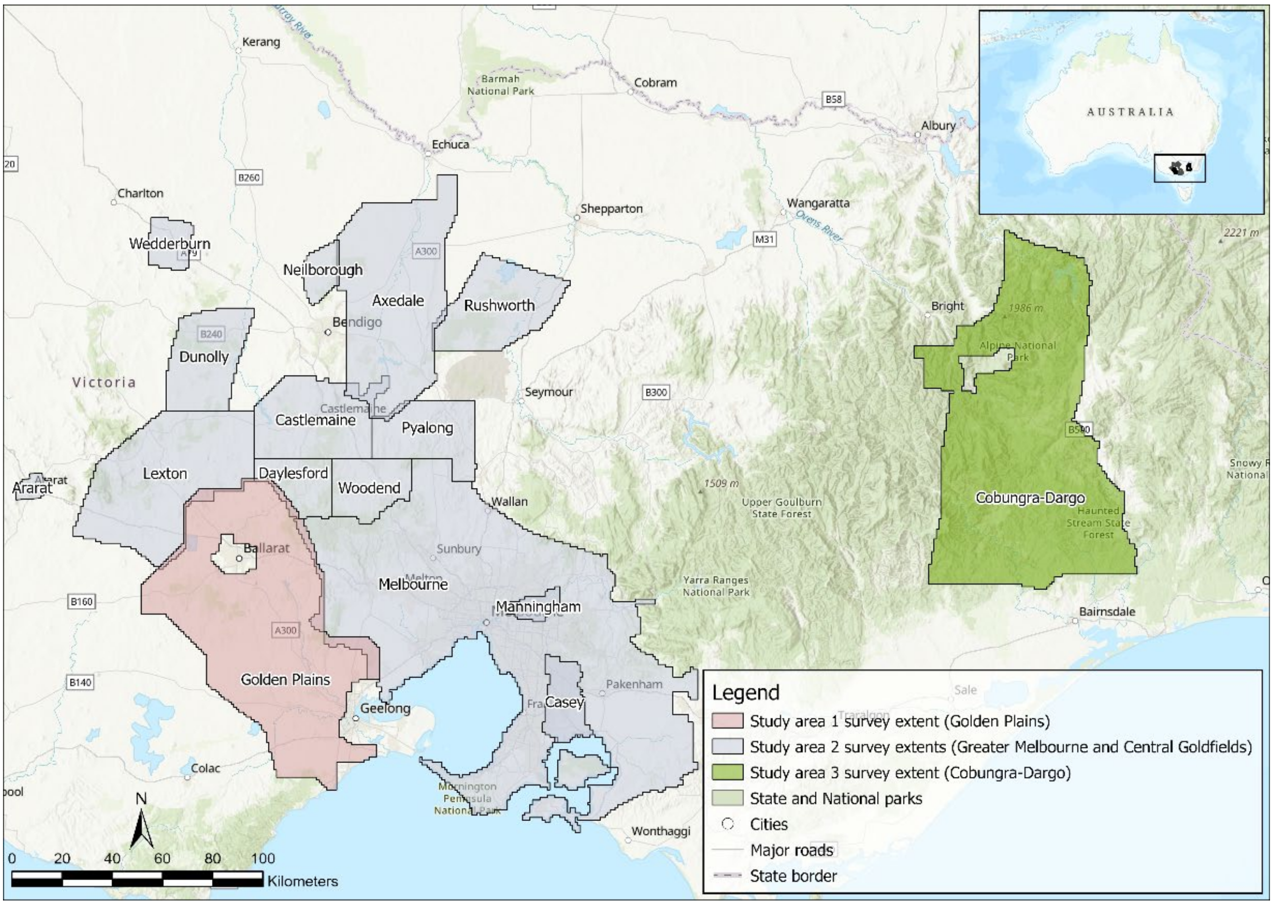

Figure 1 The location of the LiDAR surveys used in the 3 study areas

Study area 1 was defined by a single high-resolution LiDAR survey acquired in 2020. This area was selected for its extensive gold mining history, the consistent high-quality contemporary LiDAR data and flat to undulating topography. The second study area comprised 18 LiDAR surveys acquired between 2017 and 2022 and was also selected for its extensive gold mining history. The third study area was located over a newly acquired survey (2024–2025) across the Cobungra-Dargo region, which was captured for the purpose of this study.

The methods developed for the detection and classification of mining features evolved across each study area while retaining a consistent overall workflow (Figure 2).

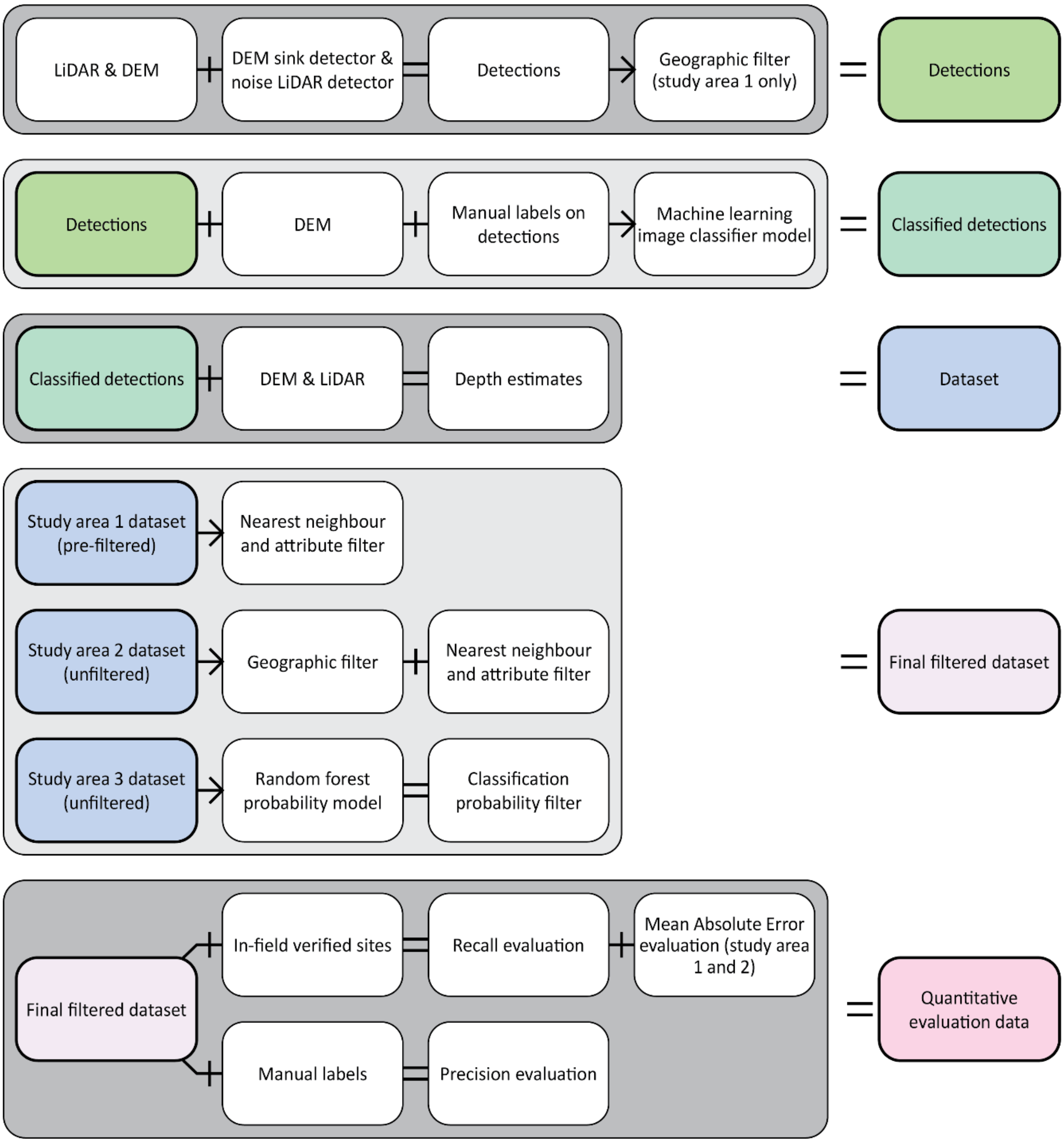

Figure 2 Workflow used for the 3 study areas, utilising LiDAR data and digital elevation models (DEMs)

Over each study area, the LiDAR point cloud was the primary input data for a noise LiDAR detector, and the DEM was used for a DEM sink detector to identify possible mine shafts and surface workings.

A subset of features identified by the DEM sink detector were manually labelled by humans using visual annotation methods to create a training dataset for the machine learning image classification model. Utilising the training data, the model then classified all detections as 'mine shaft', 'surface workings' or 'not sure/NA/other'. The DEM and the LiDAR point cloud were then used to generate depth estimates for each detection.

The detections were filtered using geographical overlays, lithological intersections and statistical analysis, and validated against observations undertaken in the field and against the manually labelled training data.

2.2 LiDAR and DEM input data

The data from 19 different LiDAR surveys were used as inputs in the machine learning classification of mining features across the 3 study areas. The Golden Plains survey over study area 1 covers 5,996 km^2, with a point density of 17 points per metre squared (pts/m^2) from which a 0.5 m DEM was generated. Study area 2, over the Greater Melbourne and Central Goldfields region, comprised 18 LiDAR surveys with point densities from 8 to 18.84 pts/m^2 and DEMs in 0.5 m and 1 m resolution. These surveys cover a total area of 25,463 km^2. The Manningham and Casey surveys overlap the Greater Melbourne survey and were preferentially used as input data because of their higher resolution. Likewise, a portion of the high-resolution Golden Plains survey from study area 1 was used where it overlapped with the western edge of the Greater Melbourne survey. The study area 3 Cobungra-Dargo LiDAR survey covers a total area of 7,635 km^2 with a point density of 36.78 pts/m^2 and a DEM resolution of 0.5 m.

Two classes from the LiDAR point cloud were the primary data inputs: ground-classified points (Class 2) and noise-classified points (Class 7), visualised in Figure 3a. Ground-classified points were used to generate DEMs (Figure 3b; grey topographic image). Due to abrupt elevation changes associated with shafts, some LiDAR points were erroneously classified as noise (Class 7) instead of ground. This error led to a smoothing effect over some mining features, resulting in false negatives. The smoothing effect is demonstrated in Figure 3b, where the mound-shaped features overlain in red noise-classified points show little to no topographical variance in their centres.

Noise-classified points, shown in red in Figure 3, were used to identify false negatives. There was some difficulty discerning the difference between true noise-classified points and misclassified noise points, though misclassified noise points often presented as clusters of multiple points, as demonstrated as a cross-section in Figure 3c. This issue primarily affected study areas 1 and 2.

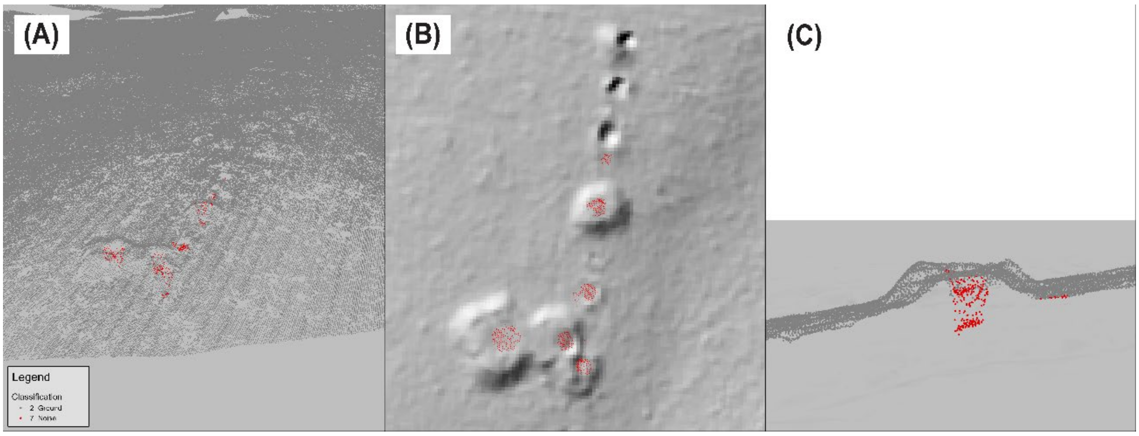

Figure 3 Visual examples of LiDAR and digital elevation models (DEM) data. (a) The LiDAR point cloud with ground-classified points (class 2, grey) and noise-classified points (class 7, red); (b) The DEM with noise-classified points overlain; (c) A cross-section of a shaft defined by noise-classified points

The LiDAR point cloud acquired for study area 3 was manually corrected, reclassifying thousands of incorrect LiDAR point classifications from noise to ground. This improved the quality of the LiDAR dataset, which generated a DEM that more accurately represented the surface topography in study area 3, resulting in more accurate detections and classifications. An additional benefit of this process was that the manual corrections could be used as proxy detections and manual labels.

2.3 Feature detection

Feature detection was required to produce datasets of potential mining features that could be labelled by humans and used to train the machine learning image classification model. Two different detectors using hydrology, topography and attribute filtering techniques were used in this study: a sink-based detector and a noise LiDAR detector.

2.3.1 Digital elevation model sink detector

Mine shafts act as hydrological sinks, allowing water to flow in but not drain until water rises above the shaft rim. As a result, sinks can be identified using a sink detection technique. The DEM sink detector applied a 'fill sink' spatial analysis function to hydrologically flood the DEMs of all 3 study areas.

The DEM sink detector algorithm involved the following steps:

1. applying a hydrological sink-filling algorithm to fill sinks on the DEM (Figures 4a, b)
2. subtracting the original DEM from the sink-filled DEM to isolate sink features (Figure 4c, light blue)
3. identifying sink edges (Figure 4c, dark blue).

Detections were then generated from the centroid locations of the sink edge polygons.

The output detections were filtered based on sink area thresholds and informed by pre-existing field observations. In study area 1, the upper area threshold was set at 30 m^2, which was increased in study area 2 to 56 m^2 following further field observations.

The minimum area threshold applied in study area 1 was 4 m^2. This was not applied to study areas 2 and 3 after mining features with surface areas less than 4 m^2 were observed in the field.

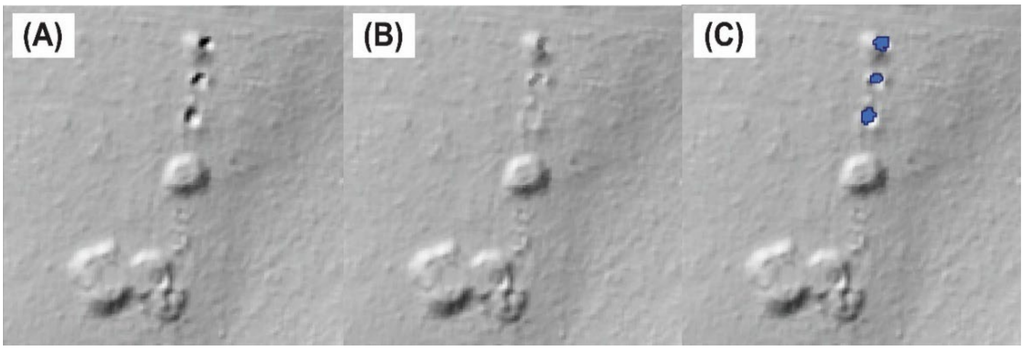

Figure 4 The steps involved to generate sink detections: (a) Original digital elevation models (DEM); (b) DEM with sinks filled to the height of adjacent cells; (c) filled sink detections (light blue) and sink edge detections (dark blue) generated after subtracting the original DEM from the filled DEM (adapted from Silver et al. 2025a, 2025b)

2.3.2 Noise LiDAR detector

The noise LiDAR detector was used to capture mining features that were missed by the DEM sink detector due to the misclassification of ground-classified points as noise-classified points in the LiDAR point cloud data. Mining features defined by noise-classified points could appear as a flat surface on the DEM and therefore do not form sinks. The application of the noise LiDAR detector resolved this issue by identifying clusters of noise-classified points below ground level. This involved the application of a 4-step algorithm:

1. filtering the LiDAR point cloud by class, retaining only noise
2. filtering by depth, retaining only points between the DEM surface and 15 m below ground level. (To capture ground-classified points or misclassified noise points below 15 m, the aircraft with the LiDAR senor would have to be positioned almost directly above the feature, which is highly unlikely. Noise-classified points below 15 m were therefore considered true noise)
3. clustering nearby points by buffering each point by a certain distance and then merging noise-classified points where their buffers overlapped
4. retaining the clusters.

This process is illustrated in Figure 5, where only 3 shafts were detected via the DEM sink detector (Figure 5a, blue circles). Application of the noise LiDAR detector identified a further 6 shafts in the area (Figures 5b, c;

red clusters of noise-classified points). These clusters highlighted potential shaft locations that appeared flat on the DEM.

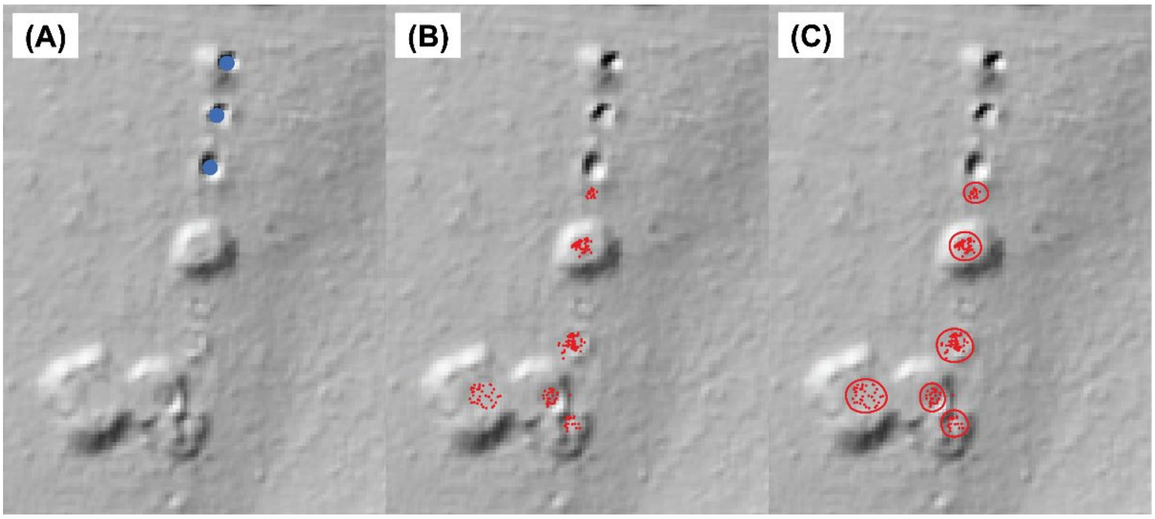

Figure 5 Ground-classified points in shafts misclassified as noise. (a) Digital elevation model (DEM) with shafts detected in blue from the sink detector; (b) DEM with noise-classified points in red; (c) Noise point clusters identified by the noise LiDAR detector (adapted from Silver et al. 2025a)

2.4 Machine learning image classification model

Datasets of manually labelled detections were used to train the machine learning image classification model. In study area 1, labelling was performed using the hillshaded DEM and a Computer Vision Annotation Tool, whereby random DEM sink detections were viewed by the manual labeller, restricted to 2 different cropped context windows (15 m and 75 m) and hillshaded at 2 different azimuths (315° and 45°) (Figure 6). This approach provided the labeller with detailed and broad views of the surrounding context and geometry of the detections. In study area 1, approximately 900 detections were labelled 'mine shaft', 'surface workings' or 'not sure/NA/other'.

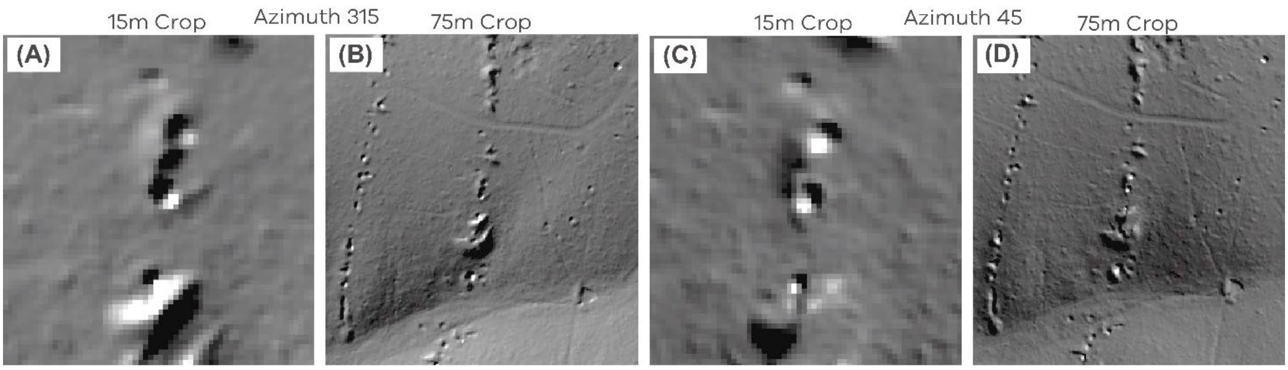

Figure 6 Example of the 2 different cropped context windows and azimuths as viewed for the manual labelling process detections in study area 1 (adapted from Silver et al. 2025a)

In study areas 2 and 3, QGIS software was used to display the DEM and all detections for manual labelling. This allowed the labeller full control of feature magnification with no restrictions on surrounding context and enabled multiple detections to be labelled at once. A targeted approach to manual, rather than random, labelling was employed to strengthen the model's ability to differentiate mining features from non-mining features.

This process was used to label approximately 5,300 detections in study area 2 and 4,500 in study area 3. The proxy detections and labels generated from the manual point cloud reclassification in study area 3 were used as an additional training dataset, comprising 8,968 detections.

70% of the training datasets were used to train the model, while 30% were used to evaluate the model (Section 3.2). This train-test split ensured that the model's performance was independently assessed, using data that was not used in the training dataset.

The model's training input data evolved with each study. Labelled detections from study area 1 were input along with a 5 m cropped context window of the raw DEM. In study area 2 the model was trained on labelled detections with 2 cropped context windows (15 m and 50 m) of hillshaded DEMs (azimuth 310° and an altitude of 45°).

The model for study area 3 was trained on labelled detections with 2 cropped context windows (15 m and 100 m) of hillshaded DEM (azimuth 310° and an altitude of 45°). Due to steep terrain in this area, 3 angled rotations of the raw DEM (0°, 120° and 240°) with a fixed light azimuth were used to provide the model with additional context to enable more accurate inferences. The model made a classification inference for each angle, which were then averaged. This process made the model results less sensitive to shade caused by terrain angle.

The parameters for each study are summarised in Table 1.

Table 1 Machine learning image classification model parameters from each study area

Parameter	Study area 1	Study area 2	Study area 3
Crop size	5 m	15 m, 50 m	15 m, 100 m
Hillshade azimuth and altitude degree	None	310°, 45°	310°, 45°
Inference method	Single inference on raw digital elevation model (DEM)	Single inference on hillshaded DEM	Inference taken as average across 3 rotations with northwest light source: 0°, 120°, 240°

The training dataset from each study area was used for each subsequent study. The training dataset from study area 1 was suitable for use in study area 2 as both regions have similarly flat to undulating landscapes. Study area 3 is located in the Victorian high country, where the terrain is steep and vastly different to study areas 1 and 2. Therefore, 3 different combinations of pre-existing and study area 3 training data were fed through the model to test the model's performance over this region. These combinations of training data were:

1. manual labels from all 3 study areas
2. only study area 3 manual labels
3. manual labels from all three study areas, including the proxy manual labels.

The test revealed that the model performed best over study area 3 when all the training datasets, including the proxy labelled detections, were used. The model then classified the input detections as 'mine shaft', 'surface workings' or 'not sure/NA/other'.

2.5 Depth estimation

In study areas 1 and 2, the depth of each classified detection was estimated using the DEM and the LiDAR point cloud, providing 2 depth estimates per detection. DEM depths were estimated by calculating the elevation difference between the highest and lowest DEM pixels. This method was expected to underestimate the depths of some shafts due to the omission of misclassified noise-classified points in the generation of the DEM.

The LiDAR depth estimation was calculated by filtering the point cloud to include only ground and noise-classified points and calculating the difference between the maximum and minimum elevation of all the ground and noise-classified points within the bounds of the detection.

The results were then matched with the original feature detection method, i.e. features detected using the DEM sink detector were attributed the depth estimated using the DEM, and similarly, detections made by the noise LiDAR detector were attributed depths estimated using the LiDAR point cloud.

All depth estimates for study area 3 were generated by calculating the difference between the maximum and minimum elevation of ground-classified points from the LiDAR point cloud (including those that had been manually reclassified).

2.6 Dataset filters

A significant number of detections were labelled 'not sure/NA/other', indicating possible false positive detections. Qualitative evaluation of the DEM and classified detections indicated that false positive detections were particularly prevalent for shallow detections and in areas of uneven ground, waterways and certain types of infrastructure. Dataset filters were used to reduce the number of false positives identified while also accounting for surrounding context, i.e. the classifications of nearby detections.

2.6.1 Geographic filter

Geographic filters were used to remove detections from locations where mining features were unlikely to occur to reduce false positive detections. For example, buffers were put in place around watercourses, the coastline and railway tracks, which produced many false positive detections due to topographical variation at these locations.

A geological filter was applied, which was generated from lithotype review of typical and atypical geological units with respect to mining activity. This review was conducted due to the high number of false positive detections generated where the You Yangs Granite outcrops, southwest of Melbourne.

Land use filters were also used to constrain detections. These included buffers around forest plantations as the land surface in these areas had a similar undulated appearance to surface workings. Likewise, buffers were applied around roads and water bodies where drains and associated infrastructure produced false positive detections.

Originally a filter was applied to remove detections from built-up areas. This was removed to ensure that real features were not excluded from towns with known mining histories, such as Ballarat and Bendigo. However, this resulted in a number of false positives, particularly across the Greater Melbourne survey. This is further discussed in Section 4.

In study area 1, this filter was applied prior to the feature detection. Experimentation with these filters required that the feature detection and machine learning image classification model methods had to be run for each experiment. To test different filters and buffer distances without the need to re-run these methods for each iteration, these filters were instead applied after the model was run in study area 2.

2.6.2 Nearest neighbour and attribute filter

The model outputs from study areas 1 and 2 were subject to a series of conditional filters. The filters applied to study area 1 were a combination of depth filters, the machine learning image classification model predictions and the model predictions of the surrounding detections. Detections were filtered and removed if:

- the LiDAR depth estimate was less than 0.8 m (the DEM sink detector should detect these features if they are true positives)
- the model prediction was 'not sure/NA/other' (unless the LiDAR depth estimate was greater than 2 m)
- more than half of the surrounding detections within a 500 m radius were predicted by the model as 'not sure/NA/other' (unless the LiDAR depth estimate was greater than 2 m).

The filter for study area 2 was far more complex. A series of tile and point cloud statistics were attributed to each detection, including the number, percentage and density of LiDAR noise and ground-classified points on each 1 km^2 tile of the LiDAR survey. These attributes were used in separate clauses for the DEM sink detections and the noise LiDAR detections. Each clause used the attached attributes to statistically filter 'not sure/NA/other' model predictions based on parameters that balanced the removal of false positive detections and the retainment of true positive detections. Establishing these parameters was an iterative process that required repeated fine tuning.

2.6.3 Probability filter

The complexities involved in creating the best possible query for study area 2 led to the use of a random forest model in study area 3. The model optimised a range of statistical parameters and assigned each classified detection a probability that the detection is a mining feature.

Like the machine learning image classification model, the random forest model was trained and tested. Both a random forest and decision tree model were trained on different groupings of manual labels and then evaluated to optimise the outputs. The optimal model was the random forest model trained on manual labels from study area 3, not including the reclassified ground points.

The output probabilities were then used to filter the final dataset.

3 Results

3.1 New LiDAR datasets

The original consolidated historical database comprised around 36,000 historical records across Victoria (Riley et al. 2025). The application of this workflow to the 3 study areas added 224,185 previously unrecorded potential mining features (Figure 7).

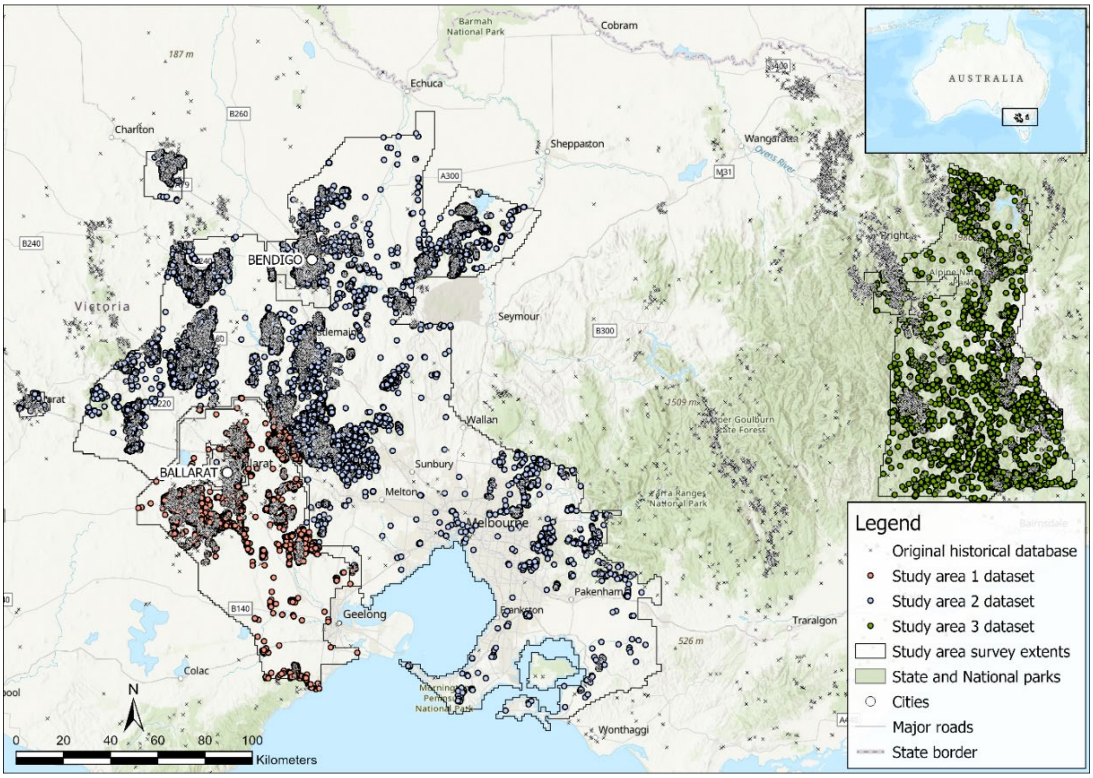

Figure 7 Distribution of the original historical database and new datasets from the study areas

Table 2 summarises the breakdown of features added to the database by feature type and detector method. To integrate these new data into the pre-existing historical database, 'mine shaft', 'surface workings' and 'not sure/NA/other' classifications are referred to as shaft, surface workings and unknown, respectively.

Table 2 The number of each feature type in each study's final dataset, including the number and percentage of features from the noise LiDAR detector method

Feature type	Study area 1 feature count		Study area 2 feature count		Study area 3 feature count	
	Digital elevation model (DEM) sink detector	Noise LiDAR detector	DEM sink detector	Noise LiDAR detector	DEM sink detector	Noise LiDAR detector
Shaft	37,578	4,020	74,090	783	18,624	767
Surface workings	31,425	0	37,793	494	420	4
Unknown	1,407	0	15,435	465	847	33
Total	70,410 (95%)	4,020 (5%)	127,318 (99%)	1,742 (1%)	19,891 (96%)	804 (4%)

Shafts account for the greatest proportion of features detected across each study area. The greatest proportion of features were identified using the DEM sink detector. Just 3% of the features detected were comprised of those from the noise LiDAR detector. The depth estimates suggested that 91% of the detections were features less than 2 m in depth, with 0.2% of features estimated to be 10 m or greater in depth.

3.2 Evaluation

The results for each study area were evaluated using 3 metrics: precision, recall and mean absolute error (MAE) of the depth estimate.

The precision metric was calculated as the proportion of machine learning mining feature classifications that matched the manual labels. It was calculated for individual classifications, e.g. shafts and, more broadly, for overall mining features (shafts or shallow workings).

$$\text{Precision} = \frac{number\ of\ manual\ mining\ feature\ labels\ matched\ by\ the\ model\ predictions}{number\ of\ manual\ mining\ feature\ labels} \tag{1}$$

Recall measured the ability of the feature detection method to identify features confirmed through field observation. Again, this could be classification specific, i.e. the number of field-observed mine shafts that were detected by the feature detection method divided by the total number of field-observed mine shafts, or, more broadly, as the proportion of all field-observed mining features also detected by the feature detection method.

$$\text{Recall} = \frac{number\ of\ field{-}verified\ mining\ features\ also\ detected\ through\ feature\ detection\ process}{number\ of\ field{-}verified\ mining\ features} \tag{2}$$

The MAE measures the average absolute difference between the predicted depth of a feature and its field-observed depth.

$$\text{Mean absolute error [depth]} = \frac{\sum |predicted\ depth - observed\ depth|}{sample\ size} \tag{3}$$

3.2.1 Precision on manually labelled detections

Precision was assessed for the whole system (machine learning image classification model and filtering process) using manually labelled detections as part of the training datasets for each study area (Table 3). The system performed well in study areas 1 and 2 but in study area 3, precision was measured at 56% when tested on manually labelled detections. This is attributed to class imbalance: in the 30% test set (1,375 detections), 94% were manually labelled as 'not sure/NA/other', leaving only 85 manually labelled mining features (6%) for precision assessment. The measure for study area 3 improved with the inclusion of the reclassified ground point test set, 64% of which were a mine shaft or surface workings.

Table 3 Precision of final filtered datasets

Precision metric	Study area 1	Study area 2	Study area 3
Mine shaft or surface workings	98%	89%	56% (manual labels only) 95% (with reclassified ground points)

3.2.2 Recall on field-observed features

To measure recall accurately, the locations of field-observed features were matched to the nearest feature detection. Table 4 outlines the results for the 3 study areas. A buffer of 10 m was applied to the in-field-observed features to account for error; however, error analysis of study area 1 indicated that the 10 m buffer was not large enough to account for a combination of GPS error and historical feature location inaccuracies. The removal of field-observed features with inaccurate locations in study area 1 improved the recall of mine shafts to 92%. Study area 2 exhibited improved recall metrics, with 89% of field-observed shafts detected.

The main sources of error in study area 3 were location inaccuracy of the field-observed features and DEM sink false negatives. Field teams were unable to undertake field work in study area 3 and so a subset of the historical database with a good level of location accuracy was used. However, the accuracy of their locations was still limited, likely by the effects of dense vegetation on GPS accuracy, and the safe distance from which the feature location could be recorded.

The reclassification of ground points in study area 3 was performed on the point cloud and therefore the feature locations were highly accurate, increasing recall.

Table 4 Recall metrics on field-observed features

Observed feature type	Study area 1	Study area 2	Study area 3
Mine shaft or surface workings	66%	88%	56% (historical field observations) 77% (with reclassified ground points)
Mine shaft only	85%	89%	–

3.2.3 Mean absolute error of estimated depths

Field-observed depths were compared to the LiDAR and DEM estimated depth attributes. Comparisons were limited to field-observed shafts that had been assigned a depth confidence of ‘certain’. While this limited the evaluation set, it more accurately reflected the MAE of the model from each study area (Table 5). The historic data used to measure recall could not be used to accurately measure the MAE for study area 3, as the confidence measures of the observed depths were approximate, or not recorded.

The MAE results indicate that the noise LiDAR depth predictions were more accurate than the DEM sink model predictions, however the sample size was much smaller.

Table 5 The mean absolute error (MAE) calculated for study areas 1 and 2

Depth estimate model	Study area 1 MAE (m)	Study area 2 MAE (m)	Study area 3 MAE (m)
Noise LiDAR model	1.22 (n=15)	0.81 (n=8)	–
Digital elevation model sink model	1.67 (n=94)	1.16 (n=47)	–

The DEM sink model tended to underestimate depth and while the noise LiDAR model produced more accurate depth estimates when discounting outliers (Figure 8), the noise LiDAR method was the sole source of outliers in the dataset. Depth predictions that were considered significant overestimations were not filtered or removed from the final dataset in order to demonstrate the limitations of the noise LiDAR model.

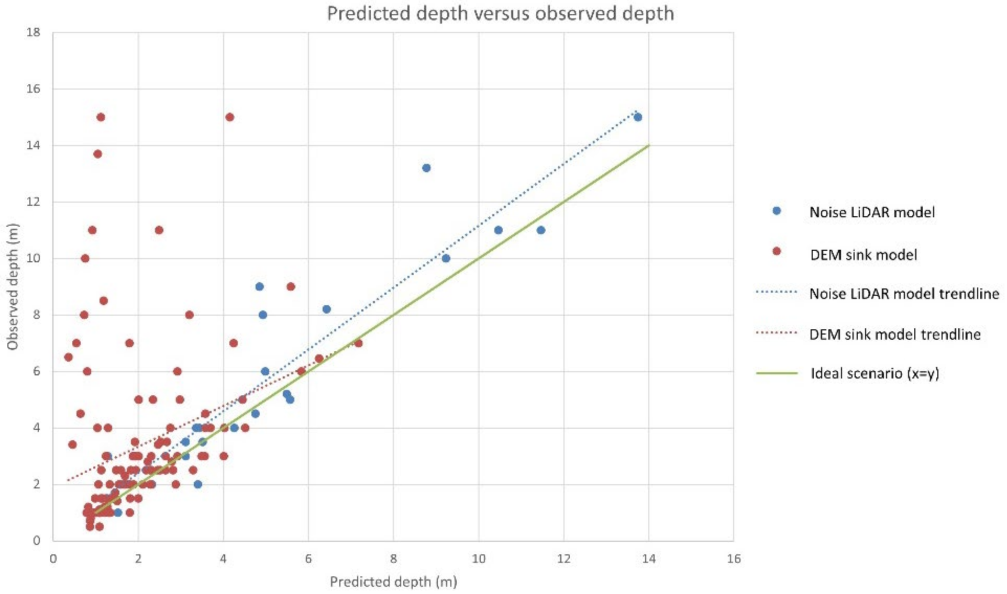

Figure 8 A plot showing the difference between the observed depth and predicted depth for the noise LiDAR and digital elevation model (DEM) sink models against the ideal scenario x = y (adapted from Silver et al. 2025a, 2025b)

4 Discussion

4.1 Machine learning image classification model

The machine learning image classification model was trained using manually labelled detections with DEM images and was therefore constrained by the quality of the underlying LiDAR data and DEM resolution. Human subjectivity was reflected in manual labelling of the training dataset. This was reduced by using random detections for manual labelling, contextualising the images for manual labelling and model training with a range of cropped context windows and shading angles, the use of geographic filters to remove reliance on visual interpretation alone and the use of evaluation methods to quantify limitations in the final datasets. Utilising the random forest probability model to filter the study area 3 dataset limited human subjectivity by optimising the statistical parameters used to refine the dataset rather than manually setting the parameters via a query.

4.2 Depth estimation

DEM-based and LiDAR-based depth estimations were limited by the accuracy and resolution of the underlying data. The DEM-based depth estimates were subject to the effects of interpolation across point cloud data when the DEM was generated, resulting in a smoothing effect across potentially true mining features. The LiDAR-based estimates relied on ground-point density and the correct classification of point cloud data. Low density data, interference below ground level (e.g. vegetation and other obstructions) and noise-classified points could have led to under- or overestimations.

4.3 Geographic filter

The filter eliminating detections from built-up areas was removed in favour of capturing potential features in towns with known mining histories, including Ballarat and Bendigo. This led to a number of false positives within the Greater Melbourne LiDAR survey area (study area 2). Of the 5,059 features detected within this survey, approximately 4,300 were in regions with well-known historical mining activity. Of the remaining features, more than 550 were false positives erroneously generated by the DEM across small data voids in the point cloud. This is a known limitation within the Greater Melbourne LiDAR survey. Other false positives were associated with urban infrastructure, including roadside drains, quarries, and disturbed ground in suburban parks and reserves, or were not captured and filtered out by the watercourse buffer.

The overall effectiveness of the geographic filter process was constrained by the accuracy and completeness of the underlying spatial datasets. Watercourse layers, land use maps and geology datasets vary in resolution and age, affecting spatial accuracy. While imperfect, the contextual use of geographic filtering reduced widespread false positives and improved the final dataset.

5 Conclusions

A large-scale machine-learning method was developed to detect and interpret previously unrecorded historical mining features using LiDAR data and digital elevation models (DEMs). First applied in the Golden Plains region of Victoria, Australia, this data-driven approach demonstrated the value of LiDAR and machine learning to substantially improve the identification of legacy mining features beyond historical records. Subsequently, the method was expanded and refined across the Greater Melbourne and Central Goldfields, and the Cobungra-Dargo, regions of Victoria. The creation of a workflow integrating LiDAR point cloud data and LiDAR-derived DEMs with geographic and geological context identified 224,185 previously unrecorded features, significantly strengthening Victoria's consolidated mining feature database. The results highlight how LiDAR provides a rapid, scalable approach for jurisdictions with historical mining legacies to better quantify and map legacy mining features.

Acknowledgements

The authors acknowledge the Traditional Owners and Custodians of the lands on which this work was conducted, including the Eastern Maar, Barengi Gadjin (Wotjobaluk, Jaadwa, Jadawadjali, Wergaia and Jupagulk peoples), Wadawurrung, Dja Dja Wurrung, Wurundjeri, Bunurong, Taungurung, Gunaikurnai and Yorta Yorta peoples. We pay our respects to their Elders, past and present, and acknowledge their continuing connection to land, waters and culture.

WSP Global Inc. was engaged to develop workflow and machine learning algorithms, and the Coordinated Imagery Program, through the Department of Transport and Planning, provided project management and oversight. Field observations were undertaken by members of the Geological Survey of Victoria and Parks Victoria.

References

Gergelova, MB, Labant, S, Mizak, J, Sustek, P & Leicher, L 2021, 'Inventory of Locations of Old Mining Works Using LiDAR Data: A Case Study in Slovakia', *Sustainability*, vol. 13, no. 12.

National Oceanic and Atmospheric Administration 2024, *What is LiDAR?*, Silver Spring, viewed 16 April 2026, https://oceanservice.noaa.gov/facts/lidar.html

Riley, C, Eid, R, Coxhead, B, Castles, K, Herley, S & Goldie Divko, L 2025, *The Former Mines and Quarries Framework Consolidated Database*, Former Mines and Quarries Framework Technical Report 1, Geological Survey of Victoria, Department of Energy, Environment and Climate Action.

Silver, ER, Dang, LH, Eid, R, Herley, SS, Riley, CP, Travers, SJ & Tran, LV 2025a, *Automated Identification of Mining Features Using Light Detection and Ranging (LiDAR) Data – Stage 1 Case Study Over the Golden Plains Region, Victoria*, Former Mines and Quarries Framework Technical Report 2, Geological Survey of Victoria, Department of Energy, Environment and Climate Action, pp. 24.

Silver, ER, Dang, LH, Eid, R, Herley, SS, Riley, CP & Tran, LV 2025b, *Automated Identification of Mining Features Using Light Detection and Ranging (LiDAR) Data – Stage 2 Greater Melbourne and Central Goldfields, Victoria*, Former Mines and Quarries Framework Technical Report 3, Geological Survey of Victoria, Department of Energy, Environment and Climate Action.

Silver, ER, Dang, LH, Herley, SS, Riley, CP, Eid, R & Tran, LV 2026, *Automated Identification of Mining Features Using Light Detection and Ranging (LiDAR) Data – Stage 3 Cobungra-Dargo Region, Victoria*, Former Mines and Quarries Framework Technical Report 5, Geological Survey of Victoria, Department of Energy, Environment and Climate Action.

U.S. Geological Survey 2025, *What is LiDAR Data and Where Can I Download it?*, Reston, viewed 16 April 2026, https://www.usgs.gov/faqs/what-lidar-data-and-where-can-i-download-it