

AI and Non-Technical Survey

Quantifying the Impact of
AI-augmented Drone Imagery Analysis

npaid.org



Norwegian
People's Aid

npaid.org

Published by Norwegian People's Aid

Author: Kyaw Lin Htut, Norwegian People's Aid

Design and layout: Magnolia design

Photo front page: Norwegian People's Aid Ukraine

Norwegian People's Aid

Postbox 88 44 Youngstorget

N-0028 Oslo

Table of Contents

Executive Summary	4
Context	5
Methodology	8
Method 1: Comparative Analysis of Trends	8
Method 2: Quasi-Experimental Research	9
Quasi-Experimental Study Input Indicators	10
Quasi Experimental Study Output Indicators	11
Quasi-Experimental Study Research Questions	12
Data Analysis and Results	13
Results of Comparative Trend Analysis (November 2023 – February 2025)	13
Summary of Data	13
Interpretation of Results	14
Main Findings of Comparative Trend Analysis	15
Results of Quasi-Experimental Study (01 August – 08 September 2025)	15
Summary of Data	15
Interpretation of Results	17
Main Findings of Quasi-Experimental Study	19
Conclusion and Limitations	20

Executive Summary

This paper examines the impact of AI-augmented drone imagery analysis on the efficiency and accuracy of Non-Technical Survey (NTS) operations conducted by Norwegian People's Aid (NPA) in Ukraine. The study focuses on the integration of AI-augmented imagery analysis, using SpotlightAI, into drone-based NTS workflows, comparing traditional manual imagery analysis with AI-assisted approaches. Using a mixed-methods design that combines longitudinal trend analysis (November 2023–February 2025) and a six-week quasi-experimental study in 2025, the research evaluates potential effects on person-hour efficiency of confirmed hazardous area (CHA) identification and evidence detection accuracy.

The findings indicate that although AI-augmented imagery analysis does not significantly increase the absolute speed of completing NTS tasks, use of AI-augmented imagery analysis corresponds positively with greater efficiency and accuracy. In the quasi-experimental study, the team using AI-augmented imagery analysis identified more than double the CHA area per person-hour compared to the team which did not use AI.

Furthermore, using the percentage in size of CHAs identified within the total area of assigned sites/villages tasked for investigation by NTS teams as a proxy for accuracy of direct evidence identification, the findings indicated that the use of AI-augmented imagery analysis produced up to 2.85 times greater NTS accuracy. Additionally, it was noted that over a one-year operational period, NTS conducted with the use of AI-augmented drone imagery analysis accounted for 76.4% of all CHA identified by NPA in Ukraine despite being used on only 13.4% of sites, suggesting a disproportionate contribution to overall land release operations.

While the paper does not establish causal mechanisms or statistically rigorous correlations, it provides empirical evidence supporting the potential value of AI integration in mine action and underscores the need for further research with larger samples and more granular activity-based time analysis.

Context

NPA has been working on the removal of landmines and other explosive ordnance (otherwise known as Land Release) in Ukraine since 2022. NPA currently has nearly 400 operational staff deployed in manual demining teams as well as making use of mine detection dogs (MDD) and remotely-operated mechanical assets. NPA also extensively uses drones in the Non-Technical survey (NTS) process, whereby hazardous areas are mapped based on direct and indirect evidence. Between 2022 and 2025, NPA's survey teams have covered over 1.1 billion square meters of land across Ukraine; of which NPA was able to identify more than 51 million square meters of Hazardous Areas (i.e. minefields, cluster munitions strikes and other battle areas). As of 31 March 2026, NPA has released nearly 24 million square meters of hazardous areas, removing nearly 1300 mines and more than 2,100 other items of explosive ordnance.

NTS is a non-intrusive process and the first step of Land Release. Survey teams will collect information from local key informants as well as identify and analyze observable evidence in the field. The main output of NTS is a Hazardous Area Report, which also includes maps of the hazardous area "polygon"- an estimated geo-referenced footprint of the minefield, cluster munitions strike or battle area. Direct evidence is recorded in the form of visually observable explosive ordnance,

debris and fragmentation from mines, explosive ordnance components, known accident locations (humans, animals, vehicles), craters and other indicators such as minefield maps produced by former combatants. Indirect evidence may be in the form of second-hand accounts of minelaying, trench lines and other indicators that military activity may have taken place. Areas with only indirect evidence may be classified as "Suspected Hazardous Areas" (SHAs) and are typically much larger and over-estimated in order to ensure a safety buffer; while areas with confirmed direct evidence are classified as "Confirmed Hazardous Areas" (CHAs) and are typically smaller in size compared to SHAs. In Ukraine, the national mine action center will assign mine action operators with "Sites" of varying size, typically corresponding with village/settlement locations of former front lines and battle areas of de-occupied territory. Operators will investigate a site and either verify that there are no explosive ordnance threats; or identify and define CHAs or SHA's if evidence of explosive ordnance contamination is present. Accurate analysis and mapping of evidence enables mine action organizations to develop "tighter", more accurate polygons of hazardous areas, thereby focusing resources to be committed to making the land safe for communities as quickly as possible. NTS is typically followed by more resource-intensive interventions in the form of Technical Survey and full Clearance. This ensures

that “all reasonable effort”¹ has been applied in accordance with National Mine Action Standards.

Whereas a typical NTS team may spend several hours or days surveying an area, a TS team may spend weeks to months conducting systematic and targeted investigations and full clearance teams may take months to years clearing a CHA. Therefore, the accuracy of hazardous area polygons has significant implications in terms of the time and resources required to release a hazardous area. The use of drones has significantly assisted NPA thus far in improving the accuracy and efficiency of NTS by allowing teams to safely (non-intrusively) observe and identify direct evidence within hazardous areas such as surface-laid mines, craters, mine fragments, ground signs; as well as indirect evidence such as trench lines. This is typically done by conducting multiple overhead flights at specified altitudes and overlapping patterns to take georeferenced optical still-imagery, which is then stitched into an “ortho-mosaic”. Prior to the introduction of AI-augmented imagery analysis, NPA staff are assigned to analyze each mosaic image frame by frame and manually identify and mark every suspected mine, item of explosive ordnance or other evidence point; the coordinates of which are entered into GIS in order to help generate hazardous area polygons and reports.

Since December 2023, NPA has been using the services of two service providers, namely Safe Pro AI’s Spotlight AI and Biodrone AS, to assist in automatically (with human in the loop) identifying and geolocating surface-laid mines, explosive ordnance and other suspected items captured from drone imagery. Both service providers provide a suite of AI assisted drone-imagery analysis tools developed for humanitarian mine action compatible with commercial off-the-shelf drones.

Once a user uploads the imagery to the web-application, the AI service automatically identifies evidence points of surface ordnance and processes 2D orthomosaic maps of the surveyed area. The AI-indicated evidence point predictions are presented to a human user on the service provider’s portal for review, which the user is then able to mark each point as confirmed, suspect or removed. Since each evidence point is georeferenced with coordinates, the evidence points can subsequently be exported into various types of GIS files including shape, KMZ and KML files as a feature layer along with an attached image of each evidence point.

For the purposes of this paper, the data captured and analysis of this data is focused on SpotlightAI. As of October 2025, NPA has processed over 1 million drone images with SpotlightAI across 750 NTS tasks in Ukraine. This has resulted in over 17,625 confirmed evidence points detected by the system, with 6,600+ hectares mapped of which 5,250 ha were classified as a CHA with a confirmed evidence point detected by SpotlightAI.

¹ <https://www.mineactionstandards.org/standards/07-11-03/>

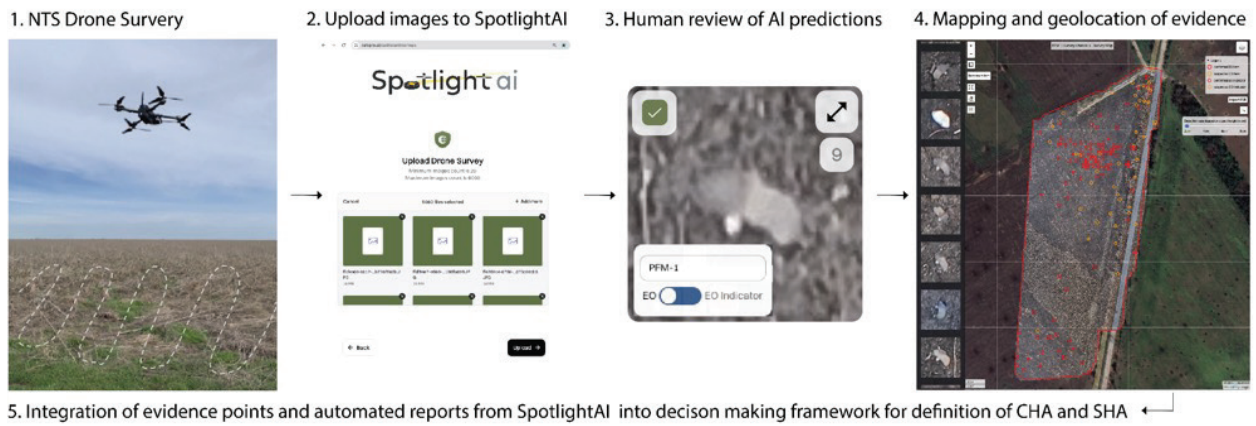


Figure 1. Illustrative workflow for how NPA integrates drones and AI-augmented imagery analysis into NTS operations using SpotlightAI. (Image courtesy of Safe Pro AI)



Figure 2. Examples of direct evidence points detected by SpotlightAI from drone imagery. The red box shows the machine learning model prediction. Detected items range from BM-21 Grad rockets, TM-62 anti-tank mines, various submunitions, and small anti-personnel mines. (Image courtesy of Safe Pro AI)

Methodology

The following methodology provides a framework to quantify and analyze the actual and potential operational gains of using AI-augmented evidence identification of drone imagery in the NTS process. In order to do so, efficiency and accuracy were measured according to the following parameters:

1. Comparing the direct cost in “person-hours” of conducting the NTS process without the use of AI-augmented evidence identification versus using AI-augmented evidence identification (i.e. SpotlightAI).
2. Extrapolating accuracy of NTS by comparing the total size of CHAs identified with AI-augmented imagery analysis versus CHAs that have been identified “manually” without AI-inputs within an operational area.

The researcher used a mixed methods approach through the following two methods, namely 1. Comparative Analysis of Trends; and 2. Quasi-Experimental Research.

Method 1: Comparative Analysis of Trends

Comparative analyses were conducted on a temporal basis between November 2023 and February 2025 of NPA's NTS operations with regards to NPA's NTS outputs (in square meters CHA identified) and AI usage outputs (in square meters processed with the assistance of AI) over the same time period in order to quantify potential effect. This trend analysis attempts to answer the following research questions:

- *Research Question 1 – Based on the Comparative Analysis of previously recorded data does use of Spotlight AI increase the total number of NTS tasks completed over the studied period?*
- *Research Question 2 – Based on the Comparative Analysis of previously recorded data, does use of Spotlight AI decrease the person-hours required per NTS task? (i.e. increase the square meters processed per person hour)*
- *Research Question 3– Based on the Comparative Analysis of previously recorded data, is the average CHA size in square meters higher after the introduction of SpotlightAI?*



Photo: Console view of NTS team conducting overhead flight. Norwegian People's Aid Ukraine

Method 2: Quasi-Experimental Research

The second part of this study was a quasi-experimental trial whereby one NTS team was assigned as a Control Group and one NTS team was assigned as an Experimental Group respectively. Each group was assigned to the same Area of Operations within a region of Ukraine², and asked to perform normal NTS duties concurrently over a 6-week period from 01 August 2025 to 08 September 2025, roughly corresponding with an operational deployment cycle. Each group was assigned different, random sites/villages to cover over the period. Since the same site or village was not investigated twice (i.e. once by each group) in a double-blind test; Method 2 is classified as quasi-experimental rather than rigorously experimental.

The variables for Method 2 can be summarized as the differences between the two groups as identified below.

- Control Group (A): This group was not given SpotlightAI to augment their NTS process.
- Experimental Group (B): This group was given SpotlightAI to augment their NTS process.

Therefore, indicators denoted “A” pertain to the Control Group, and indicators denoted “B” pertain to the Experimental Group. The Inputs (in person hours) and Outputs (in square meters and items identified) were measured of both groups. The Input and Output indicators are used to assess efficiency on a person-hour basis; while accuracy is assessed by comparing the ratio of various types of outputs.

² Due to security concerns, the author has chosen not to publicly disclose the region in which NPA is assigned its Area of Operations.

Quasi-Experimental Study Input Indicators

The person-hours spent by staff in each group; namely the Control Group and Experimental Group were measured as Input Indicators. Figure 3 is an illustrative breakdown of Input Indicators

(Indicators 1 to 3) and sub-indicators contributing to person-hours which were quantitatively compared between Control Group (A) and Experimental Group (B) only. As a limitation of the study, reference ranges for inputs from other teams were not recorded.

Sub-processes Indicators	Process steps and working techniques	NTS Team (A)			NTS Team (B)			
		# Staff	Hrs	Total Person Hrs	# Staff	Hrs	Total Person Hrs	
Indicator 1. Desk study on evidence (Input)	1.1. Analysis of available information on evidence							
	1.2. Description and classification of evidence							
Input Indicator 2. NTS Fieldwork (Input)	2.1. Analysis of characteristic of NTS task							
	2.2. Analysis of missing data and task planning including drone flight							
	2.3. Field data gathering	2.3.1 Transport from office to tasks						
		2.3.2 Measurements for positioning of evidence on the ground						
		2.3.3 Collection of information from key information interviews						
		2.3.4 Drone flying to collect data						
		2.3.5 On-site data processing and reporting (on survey, on drone flight)						
	2.4. Sub division of NTS task							
	2.5. Impact assessment for affected community							
	2.6. Priority setting for land release							
2.7. Cancellation of SHAs (if any)								
Input Indicator 3. Reporting and recording (Input)	3.1. Analysis and processing of drone imagery							
	3.2. Analysis and processing of other recorded information							
	3.3. Reporting and recording of hazardous area report							
	3.4. Quality assurance checks in database							
	3.5. Compiling reports for national authorities							
	3.6 Time spent on Other Activities							
	Total							

Figure 3. Indicators of Quasi-Experimental NTS Processes (Group A/B) - Illustrative Template

Quasi Experimental Study Output Indicators

The following outputs were measured in relation to the Control Group (A) and Experimental Group (B). By the time this study was conducted, all NPA NTS teams have been using AI-augmented analysis of drone imagery during the Experimental Period, except for the Control Group (A).

- *Indicator 4: NTS Tasks (CHA/SHA) Completed in the Quasi-Experimental period. This indicator serves as a proxy through which the speed of surveying NTS tasks can be quantified by the ratio of square meters of NTS Tasks processed in person hours; as well as the average total hours it takes to complete the identification of a CHA or SHA.*

The underlying assumption is that the Group with that can process more NTS tasks per person hour, and complete identification of a CHA and SHA in less time is the faster working Group. These indices are reported by the following sub-indicators.

- Indicator 4.1 Total NTS tasks (CHA/SHA) completed by the Groups (A, B).
- Indicator 4.2. Average time to complete CHA identification per NTS task (person hours) by the Groups (A,B)
- Indicator 4.3. Average time to complete SHA identification of NTS task (person hours) by the Groups (A, B)
- Indicator 4.4. Total Number of Sites/Villages³ covered during NTS by the Groups (A, B)
- Indicator 4.5. Total Area of Sites/Villages covered to the Groups (A,B) for NTS

- *Indicator 5: Measurement of average SHA and CHA polygon size identified per Person-Hour in the Quasi-Experimental period. This indicator serves as a proxy through which efficiency of surveying NTS tasks can be quantified by ratio of square meters processed in person hours.*

The underlying assumption is that the Group with more CHA area identified per hour is the more efficient Group. These indices are reported by the following sub-indicators.

- Indicator 5.1 Total SHA Area in square meters identified by the Groups (A, B).
- Indicator 5.2 Total CHA Area in square meters identified by the Groups (A, B).
- Indicator 5.3 Average number of square meters of SHA per person-hour identified by the Groups (A, B).
- Indicator 5.4 Average number of square meters of CHA identified per person-hour by the Groups (A, B).

- *Indicator 6: Measurement of CHA identification accuracy. This indicator serves as a proxy through which the accuracy of identifying CHAs can be quantified by comparing the percentage of CHA square meters in relation to the total size of assigned sites/villages; as well as CHAs compared to SHAs. It is important to note that only the identification of CHAs is considered as being “accurate NTS”; and not of SHAs. This is because CHAs would be indicative of finding direct evidence.*

³ Note: In Ukraine, sites (typically tied to a village) are assigned by the National Mine Action Center for mine action operators to investigate with their NTS teams. NTS teams move from one site to another to identify direct and indirect evidence. When no CHAs or SHAs are identified due to lack of evidence, the NTS team does not record any hazardous area and the site is recommended to be removed from the national database. If there is initial evidence to indicate the presence of a hazardous area, the NTS team will conduct further non-intrusive investigation and definitively define the extent of the hazardous area polygon (as either CHA or SHA) and compile a hazardous area report. It may be possible that evidence may warrant an EOD callout at a site without the need for defining a hazardous area after a comprehensive assessment is made by the NTS team. For the purposes of this Quasi-Experimental study, sites were assigned randomly to each NTS team within NPA's area of responsibility.

The underlying assumption is that the Group with a higher percentage of CHAs identified compared to total area of the assigned sites/villages has greater NTS accuracy. These indices are reported by the following sub-indicators.

- Indicator 6.1 Percentage of CHA square meters identified compared to total area of Sites/ villages assigned in square meters recorded by the Groups (A, B).
- Indicator 6.2 Proportion of CHA square meters identified compared to SHA in square meters recorded by the Groups (A, B).
- *Indicator 7: Measurement of evidence point/ item identification in EOD Callouts. This indicator serves as a proxy through which the accuracy of identifying individual items of explosive ordnance can be quantified.*

The underlying assumption is that the Group with a higher number of average items identified per callout has more item identification accuracy. This is reported by the following indicator:

- Indicator 7.1 Average number of items identified per EOD callout recorded by the Groups (A, B).

Quasi-Experimental Study Research Questions

Taking into consideration the Input and Output indicators, the Quasi-Experimental study attempts to answer the following research questions:

- *Research Question 4 – Based on the Quasi-Experimental Study of recorded input and output data during the set period, does the use of AI-augmented imagery analysis increase the speed at which NTS tasks are conducted?*
- *Research Question 5 – Based on the Quasi-Experimental Study of recorded data during the set period, does use of AI-augmented imagery analysis increase the efficiency through which CHAs are identified?*
- *Research Question 6 – Based on the Quasi-Experimental Study of recorded data during the set period, is the accuracy of NTS, measured as a percentage of CHA size to assigned Site/Village size in square meters greater with the introduction of AI-augmented imagery analysis?*
- *Research Question 7 – Based on the Quasi-Experimental Study of recorded data during the set period, is the accuracy of NTS, measured as a proportion of CHA to SHA size in square meters higher with the introduction of AI-augmented imagery analysis?*
- *Research Question 8 – Based on the Quasi-Experimental Study of recorded data during the set period, is the accuracy of explosive ordnance item identification, measured as the average number of items identified per EOD callout, higher with the introduction of AI-augmented imagery analysis?*

Data Analysis and Results

Results of Comparative Trend Analysis (November 2023 – February 2025)

Summary of Data

The Comparative Trend Analysis compares the outputs per month of CHA identified over the period of November 2023 to February 2025. This is done by measuring the following factors:

- Total number of NTS sites surveyed by NPA in Ukraine
- Total number of NTS sites surveyed by NPA in Ukraine using SpotlightAI

- Total size of CHAs identified by NPA in Ukraine (in hectares)
- Total size of areas identified by SpotlightAI as probable CHA (in hectares)

This data is represented in the subsequent charts, with AI usage data provided courtesy of Safe Pro AI.

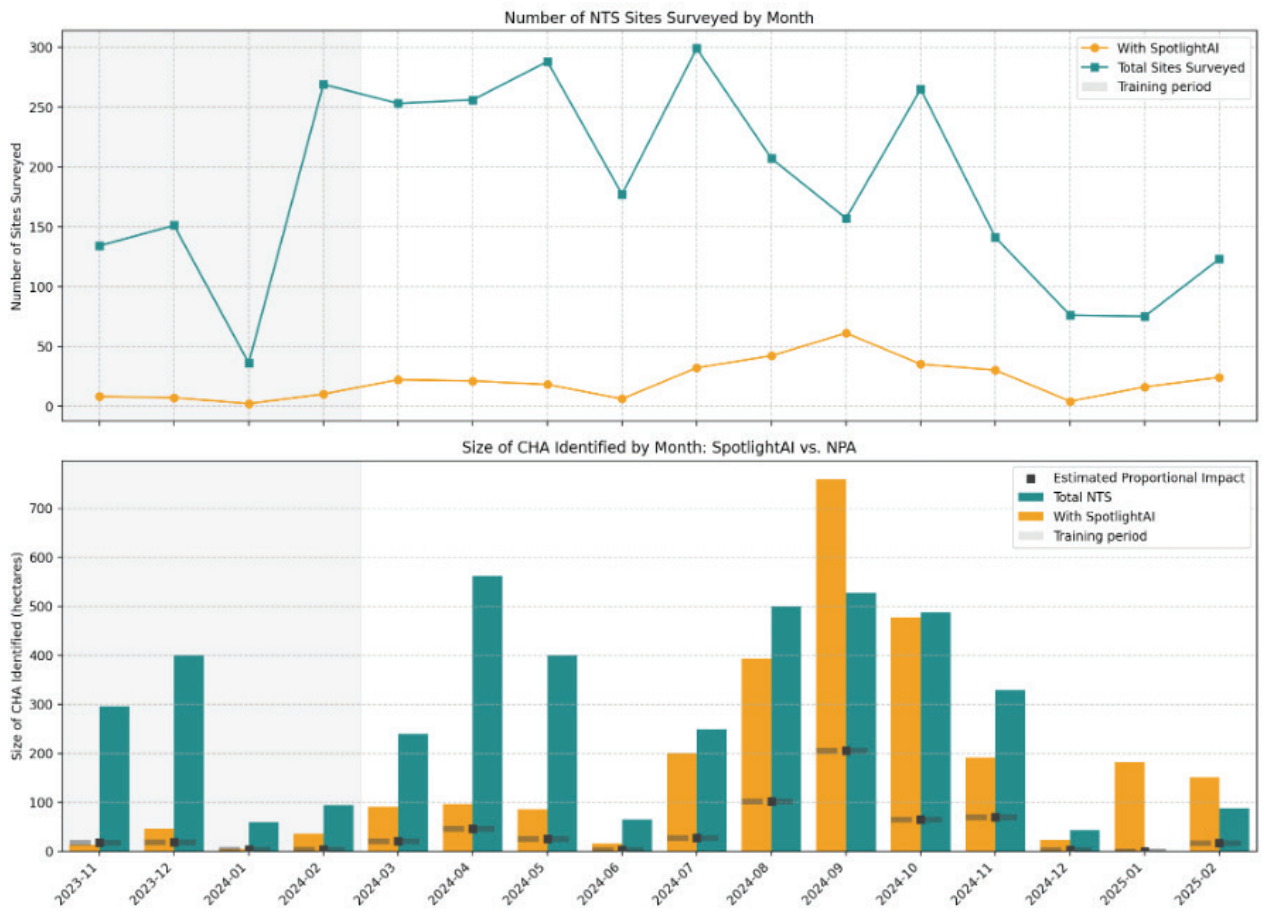


Figure 4: Monthly NTS outputs Nov 2023–Feb 2025.

The line graphs in the top panel in the figure above shows total NPA sites surveyed (teal) versus sites surveyed with SpotlightAI (orange). The grey shaded region (Nov 2023–Feb 2024) indicates AI training period.

The bar graphs in the bottom panel shows the size of CHA identified (in hectares) by total NTS (teal) and by SpotlightAI (gold). Dark square markers and dashed horizontal segments depict the “estimated proportional effect” (SpotlightAI share of sites × total CHA) – a baseline expectation if CHA discovery scaled only with site share. SpotlightAI usage bars exceeding these markers indicate higher CHA identification efficiency relative to surveyed site share, especially after training completion.

Interpretation of Results

Based on these factors, we are able to establish the usage of SpotlightAI out of all NTS tasks that were surveyed by NPA. From March 2024–2025, the percentage of sites surveyed with SpotlightAI was 13.4%. The bottom panel shows the size of CHA in hectares defined per month across all NTS in Ukraine (teal bars) and the NTS using SpotlightAI. Over the course of 1 year, AI-augmented imagery analysis was used in some way to assist in identifying or processing 76.4% of all CHA measured in total size in square meters, despite only being used for 13.4% of NTS tasks.

It must be taken into consideration that this Comparative Analysis was limited to comparing only outputs. A major limitation of this portion of the study was that there was insufficient data

collected on NPA's person-hour inputs over this period, and it was not possible to do so retroactively. Therefore, the author could not make a comparison of person-hours expended per output (i.e. before and after the use of Spotlight AI). It also does not take into account duplication of activities: for example, an NTS team using Spotlight AI to help identify a CHA; and a clearance team returning to the same CHA to conduct a pre-clearance assessment using Spotlight AI (which would still be indicated as CHA area in Spotlight's system).

This caveat can potentially explain instances in the data where more CHA square meters were indicated by Spotlight AI in September 2024 than total NTS conducted by NPA for that month. Additionally, for months like January and February 2025, due to weather restrictions and snow the NTS teams were unable to conduct new drone surveys, and instead used this time to analyze a backlog of old surveys using SpotlightAI. Nevertheless, the overall trend from the analysis shows that utilizing SpotlightAI has a disproportionate contribution for defining CHA in NTS activities.

Main Findings of Comparative Trend Analysis

The main findings of the trend analysis answers Research Questions 1 to 3 as follows:

- *Research Question 1* – Based on the Comparative Analysis of previously recorded data, it was not possible to ascertain whether the use of Spotlight AI increased the total number of NTS tasks completed over the studied period. The trend of Spotlight AI use on all NTS tasks increased as the total number of NTS tasks increased; but it is inconclusive to say whether or not this was a contributing factor to completing more NTS tasks.
- *Research Question 2* – Based on the Comparative Analysis of previously recorded data, it was not possible to ascertain whether Spotlight AI decreased the person-hours required per NTS task. Person-hours for staff working on discrete activities were not recorded during this period.
- *Research Question 3* – Based on the Comparative Analysis of previously recorded data, the average CHA size in square meters/hectares was higher after the introduction of SpotlightAI.

Results of Quasi-Experimental Study (01 August – 08 September 2025)

Summary of Data

Over the course of the Quasi-Experimental period, the total of all Input indicators (Indicators 1 to 3) can be summarized for each respective Group as follows:

- Control Group A (NTS Team not provided with AI-augmented imagery analysis) spent 323.5 total person-hours working on NTS tasks and EOD callouts over the experimental period. This team worked on 11 randomly assigned sites/villages during this period, of which 2 CHAs were identified.
- Experimental Group B (NTS Team provided with AI-augmented imagery analysis) spent 326 total person-hours working on NTS tasks and EOD callouts over the experimental period. This team worked on 4 randomly assigned sites/villages during this period, of which 2 CHAs were identified.

A breakdown of person-hours input by discrete activity is provided in the following table.

Sub-processes Indicators	Process steps and working techniques	NTS Team (A)			NTS Team (B)			
		# Staff	Hrs	Total Person Hrs	# Staff	Hrs	Total Person Hrs	
Indicator 1. Desk study on evidence (Input)	1.1. Analysis of available information on evidence	2	4.75	9.5	2	8	16	
	1.2. Description and classification of evidence	2	1	2	2	5.5	11	
Indicator 2. NTS Fieldwork (Input)	2.1. Analysis of characteristic of NTS task	2	7	14	2	11.75	23.5	
	2.2. Analysis of missing data and task planning including drone flight	2	4.25	8.5	2	4.75	9.5	
	2.3. Field data gathering	2.3.1 Transport from office to tasks	2	23.25	46.5	2	20.5	41
		2.3.2 Measurements for positioning of evidence on the ground	2	2	4	2	3.5	7
		2.3.3 Collection of information from key information interviews	2	3	6	2	4.25	8.5
		2.3.4 Drone flying to collect data	2	6.75	13.5	2	13.5	27
		2.3.5 On-site data processing and reporting (on survey, on drone flight)	2	0	0	2	0.25	0.5
	2.4. Sub division of NTS task	2	0	0	2	0	0	
	2.5. Impact assessment for affected community	2	0	0	2	0.5	1	
	2.6. Priority setting for land release	2	0	0	2	0.5	1	
2.7. Cancellation of SHAs (if any)	2	0	0	2	0	0		
Indicator 3. Reporting and recording (Input)	3.1. Analysis and processing of drone imagery	2	22.25	44.5	2	23	46	
	3.2. Analysis and processing of other recorded information	2	5.25	10.5	2	9.25	18.5	
	3.3. Reporting and recording of hazardous area report	2	20.25	40.5	2	17.75	35.5	
	3.4. Quality assurance checks in database	2	0	0	2	0	0	
	3.5. Compiling reports for national authorities	2	0.25	0.5	2	1	2	
	3.6 Time spent on Other Activities	2	61.75	123.5	2	39	78	
	Total			323.5			326	

Figure 5: Input Indicators of Quasi-Experimental Study

Over the course of Quasi-Experimental period, the following Output Indicators (4 to 7) can be

summarized for each respective Group as follows in the table below:

Indicator	Control Group (A)	Experimental Group (B)
Indicator 4.1 Total NTS Tasks (CHA/SHA) Completed	2	2
Indicator 4.2. Average time to complete CHA identification per NTS task (person hours)	161.75 hours	163 hours
Indicator 4.3. Average time to complete SHA identification of NTS task (person hours)	N/A	N/A
Indicator 4.4. Total Number of Sites/Villages covered during NTS	11	4
Indicator 4.5 Total Area of the Sites/Villages covered	1,090,000	800,400
Indicator 5.1 Total SHA Area in square meters identified	N/A	N/A
Indicator 5.2 Total CHA Area in square meters identified by the Groups	18655	
	39072	
Indicator 5.3 Average number of square meters of SHA per person hour identified	N/A	N/A
Indicator 5.4 Average number of square meters of CHA identified per person-hour	57.66	119.85
Indicator 6.1 Percentage of CHA square meters identified compared to assigned Sites/Villages in square meters	1.71%	4.88%
Indicator 6.2 Proportion of CHA square meters identified compared to SHA in square meters	N/A	N/A
Indicator 7.1 Average number of items identified per EOD callout	1.77 (23 items over 13 callouts)	7.75 (31 items over 4 callouts)

Figure 6. Output Indicators of Quasi-Experimental Study

Interpretation of Results

The results of the quasi-experimental study are interpreted in accordance with the research questions below.

- *Research Question 4 –Based on the Quasi-Experimental Study of recorded input and output data during the set period, does the use of AI-augmented imagery analysis increase the speed at which NTS tasks are conducted?*
- The Experimental Group spent a total of 326 person-hours in the study period, while the

Control Group spent a total of 323.5 person-hours. Both groups were able to identify two CHAs each in this period. This equated to 161.75 hours per CHA identified for the Control Group and 163 hours for the Experimental Group.

- The Experimental Group covered 4 randomly assigned sites/villages with a total area of 800,400m² and the Control Group covered 11 randomly assigned sites/villages with a total area of 1,090,000.
- The results suggest that the difference in person hours expended per CHA identified is negligible. Therefore, when comparing Indicator 4.1 to Indicators 4.2, 4.3 and 4.4, the

use of AI-augmented imagery analysis does not appear to increase the absolute speed at which individual NTS tasks or coverage of sites/villages is completed.

- *Research Question 5– Based on the Quasi-Experimental Study of recorded data during the set period, does use of AI-augmented imagery analysis increase the efficiency through which CHAs are identified?*
- No SHAs were identified by either Group in the study period.
- The Experimental Group identified 39,072m² of CHA while the Control Group identified 18,655m² of CHA.
- This equated to 119.85m² of CHA identified per person hour expended by the Experimental Group and 57.66m² of CHA identified per person hour expended by the Control Group.
- Using CHA size in square meters identified per person hour as a proxy for efficiency, the use of AI-augmented imagery analysis appears to have an effect on increasing efficiency through which CHAs are identified.
- *Research Question 6– Based on the Quasi-Experimental Study of recorded data during the set period, is the accuracy of NTS, measured as a percentage of CHA area to assigned Site/Village size in square meters greater with the introduction of AI-augmented imagery analysis?*
- The Experimental Group identified 39,072m² of CHA over a total site/village area of 800,400m². This equates to identifying CHAs on 4.88% of the covered site/village area.
- The Control Group identified 18,655m² of CHA over a total site/village area of 1,090,000m². This equates to identifying CHAs on 1.71% of the covered site/village area.
- Although the Experimental Group covered fewer sites/villages than the Control Group (4 to 11 respectively), it is also worth noting that the sites/villages encountered by the Experimental Group were larger in size (200,100m² on average for the Experimental Group versus 99,091m² for the Control Group).
- Using the percentage of CHA identified (i.e. proxy for the indication of direct evidence) per total site/village size as a proxy for accuracy of the NTS process, the introduction of AI-augmented imagery analysis appears to have a positive effect on NTS accuracy based on the data presented.
- *Research Question 7– Based on the Quasi-Experimental Study of recorded data during the set period, is the accuracy of NTS, measured as a proportion of CHA to SHA size in square meters higher with the introduction of AI-augmented imagery analysis?*
- Since neither Group identified any SHAs during the study period, this research question was not applicable for this study.
- *Research Question 8– Based on the Quasi-Experimental Study of recorded data during the set period, is the accuracy of explosive ordnance item identification, measured as the average number of items identified per EOD callout, higher with the introduction of AI-augmented imagery analysis?*
- The Experimental Group identified 31 items of explosive ordnance over the course of 4 callouts. This averaged to identifying 7.75 items per EOD callout.
- The Control Group identified 23 items of explosive ordnance over the course of 13 callouts. This averaged to identifying 1.77 items per EOD callout.
- Based on the higher average number of items identified by the Experimental Group, the data suggests that the introduction of AI-augmented imagery analysis may have a positive effect on the accuracy of explosive ordnance item identification.

Main Findings of Quasi-Experimental Study

The data in support of answering the quasi-experimental study's Research Questions (4 to 8) indicate the following findings:

- For the same number of person-hours spent on conducting NTS, the Experimental Group (B) was able to analyze 2.08 more times larger CHA size compared to the Control Group (A). In other words, this equates to a **108% greater processing of CHAs square meters per person hour for NTS tasks by the Experimental Group.**
- The Experimental Group (B) identified CHAs in 4.88% of the total site/village area assigned, while the Control Group (A) identified CHAs in 1.71% of the total site/village area assigned.

Using this metric as a proxy for accuracy in CHA identification (indication of direct evidence), **the Experimental Group (B) was 2.85 times (185%) more accurate in their NTS processes of identifying CHAs** compared to the Control Group.

- As an additional observation, the total CHA size identified by both teams in the study period was 57,727m². It was noted that the Experimental Group's outputs constituted 68% of CHA area identified, while the Control Group's outputs constituted 32% of CHA area identified.

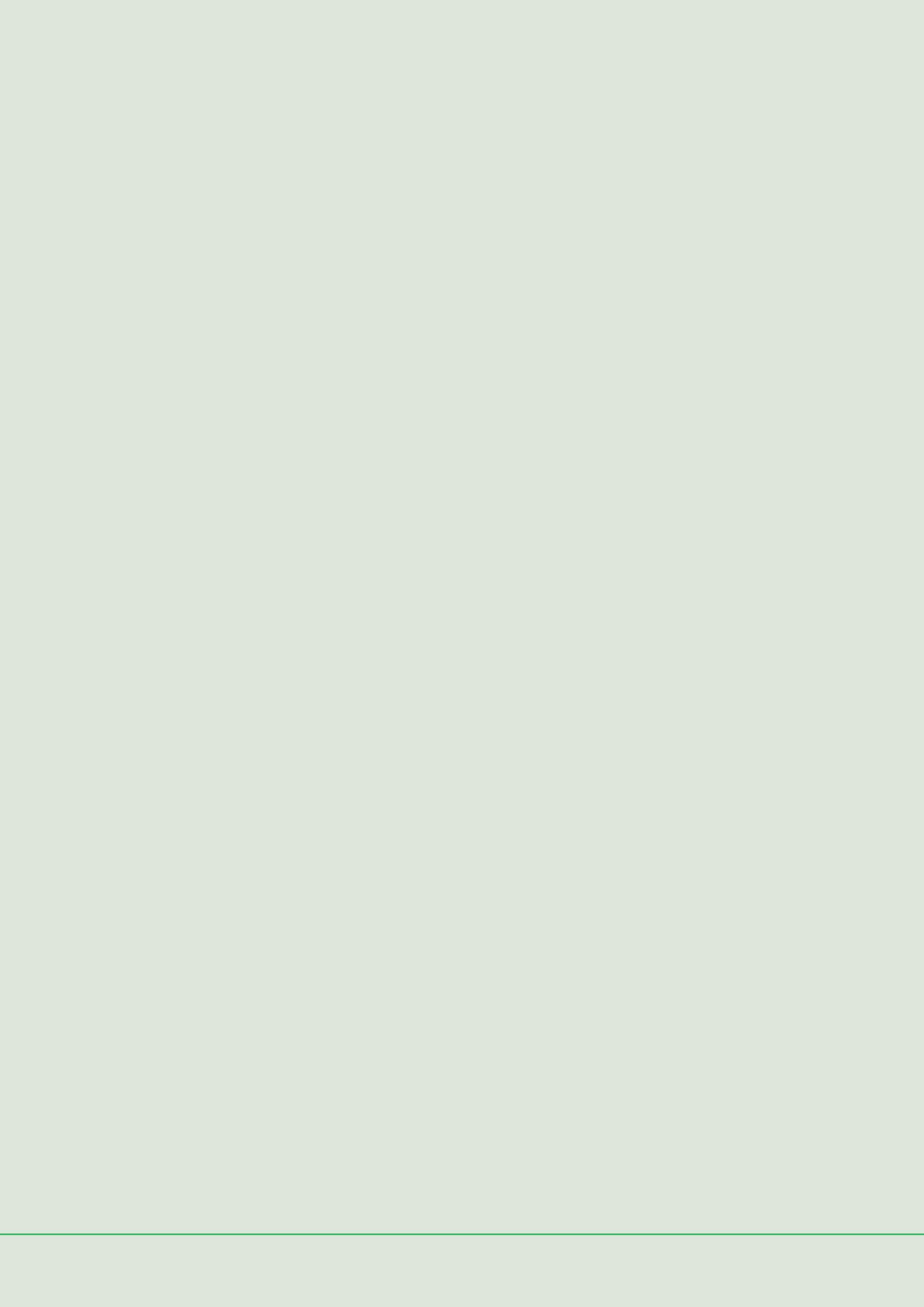
Conclusion and Limitations

The findings of the quasi-experimental study, analyzed primarily through the use of comparative ratios, appears to support a hypothesis that the use of AI-augmented imagery analysis has a positive effect on the efficiency and accuracy of the NTS processes (and supplementary EOD callout processes). The data does not appear to support a hypothesis that use of AI-augmented imagery analysis has a positive effect on the speed at which NTS is completed (as measured by the total number of NTS tasks completed over a given period), since both Groups identified the same number of CHAs over the study period expending roughly similar total person hours. However, it should be noted that the Experimental Group identified larger CHAs compared to the Control Group.

Analyzing the cumulative use of AI-augmented imagery analysis services compared to NPA's total NTS outputs over an initial 1-year period appear to suggest a positive relationship between service use and increased NTS outputs. Although AI-augmented imagery analysis services were only used on 13.4% of tasked NTS activities, these constituted 76.4% of all CHA identified by NPA

during this period. This appears to suggest a disproportionate impact in the use of AI-augmented imagery analysis by a ratio of up to 5.7.

However, it must be noted that a major limitation of this study is that it does not attempt to establish causal relationships and mechanisms through which the corresponding phenomenon occur. It must be acknowledged that the quasi-experimental study measures the total time in person-hours spent by the Groups, and does not analyze and compare in more granular detail the person hours spent on discrete activities resulting in specific NTS outputs. Future studies which aim to replicate this paper's findings and build upon its methodology should be encouraged to include a breakdown of discrete time-based activities in further detail in order to establish statistical correlation and causality between specific NTS outputs. Larger sample sizes should also be considered when conducting future statistical analyses and experimental/quasi-experimental studies.



npaid.org

Postbox8844 Youngstorget
N-0028 Oslo



Norwegian
People's Aid