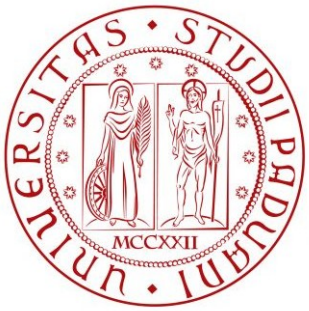


UNIVERSITÀ DEGLI STUDI DI PADOVA

Administrative unit: **University of Padova**

Department: **Land, Environment, Agriculture and Forestry (TESAF)**

PhD Program: **Land, Environment, Resources and Health (LERH)**

Batch: XXXVIII

REMOTE SENSING–BASED TERRAIN ANALYSIS FOR TIMBER HARVESTING PLANNING IN DIVERSE FOREST ENVIRONMENTS

The thesis was financially supported by the ETN Skill-For-Action project (Marie Skłodowska-Curie, Grant Agreement No. 956355) and by the Agritech National Research Center (PNRR/NextGenerationEU, Grant Agreement No. CN00000022).

PhD Program Coordinator: Prof. Marco Borga

Supervisors: Prof. Stefano Grigolato, Assoc. Prof. Bruce Talbot

Co-Supervisor: Assoc. Prof. Francesco Pirotti

PhD candidate: Gunta Grube

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

Remote Sensing–Based Terrain Analysis for Timber Harvesting Planning in Diverse Forest Environments

by

Gunta Grube

Dissertation presented in fulfilment of the requirements for the degree Doctor of Philosophy in the Department of Forestry and Wood Science, Faculty of AgriSciences, Stellenbosch University.

Supervisors: Prof. Stefano Grigolato
Assoc. Prof. Bruce Talbot
Co-supervisor: Assoc. Prof. Francesco Pirotti

January 2026

I dedicate this dissertation to my family.

Acknowledgements

First and foremost, I would like to express my heartfelt thanks to my supervisors, Prof. Stefano Grigolato and Assoc. Prof. Bruce Talbot, as well as my co-supervisor, Assoc. Prof. Francesco Pirotti. Their guidance, encouragement and support throughout this journey have been invaluable. I could not have asked for better mentors and a more supportive environment. I believe that we grow in the right place, with the right people, and I am profoundly grateful to have experienced this throughout my doctoral research.

I want to express my deepest gratitude to my mother, my sister, my brother, and the rest of my family for always being there for me and for one another. I am deeply grateful to have you in my life.

My warmest thanks go to all my Skill4Action fellows for sharing this extraordinary journey with me. Being part of such an inspiring project has been an incredible experience. I am equally grateful to my dear colleagues for their support, encouragement, and friendship throughout this time.

I would also like to thank everyone who helped me in any way throughout this journey, including the administrative and research staff whose assistance made countless tasks smoother. I am thankful to the Department of Land, Environment, Agriculture and Forestry (TESAF) at the University of Padova for providing the resources and academic environment that enabled this research. My sincere thanks also go to Stellenbosch University for hosting me during my research stay and for providing such a supportive and encouraging environment. I also wish to thank Anton Kunneke for his valuable assistance with data collection and fieldwork.

I wish to acknowledge the financial support provided by the Marie Skłodowska-Curie Actions Innovative Training Network (MSCA-ITN) 'Skill4Action' of the European Union's Horizon 2020 research and innovation programme, and the Piano Nazionale di Ripresa e Resilienza (PNRR), Italy's National Recovery and Resilience Plan, as part of the European Next Generation EU programme. Their support has been essential for the successful completion of my doctoral research.

Lastly, I would like to thank myself for believing, persevering, and showing up through every challenge - for working hard regardless of the outcome, for striving to do more right than wrong, and for staying true to who I am throughout this journey.

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 07 January 2026 . . . . . .

ABSTRACT

Remote Sensing–Based Terrain Analysis for Timber Harvesting Planning in Diverse Forest Environments

Sustainable mechanised timber harvesting depends on accurate terrain information to balance operational efficiency, cost-effectiveness, machine safety, and environmental protection. Conventional terrain classification systems (TCS) used in forestry are typically based on stand-level descriptions, manual field surveys, and expert judgement. While effective for broad planning purposes, these approaches provide limited spatial detail, are prone to observer bias, and are challenging to scale across large forest landscapes.

This dissertation investigates the potential of remote sensing-based terrain analysis to quantify terrain attributes relevant for mechanised harvesting planning and to evaluate their analytical reliability and operational feasibility across different forest environments, sensor platforms, and spatial scales. The research is structured around three core terrain components underpinning established TCS frameworks: soil conditions, surface roughness, and slope.

Using case studies based on UAV Structure-from-Motion (SfM) photogrammetry, airborne laser scanning (ALS), UAV LiDAR, and handheld and terrestrial laser scanning (HLS, TLS), terrain metrics were derived and analysed in relation to (a) machine-induced soil deformation, (b) near-ground obstacles, and (c) steepness and terrain complexity across spatial scales.

For soil conditions, the work examined whether stump proximity and stump/root reinforcement proxies explain within-site variability in rut development under repeated machine traffic and evaluated models for rut-depth prediction. Models combining terrain-derived predictors with operational variables achieved $R^2 = 0.69–0.85$ when predicting observed rut depth across machine passes and trail configurations.

For surface roughness, the research operationalises the obstacle-based roughness concept used in forestry terrain classification by combining UAV photogrammetry with deep-learning-based obstacle segmentation. In the high-visibility study setting, this achieved a 95.6% obstacle detection rate (86/90) relative to field-marked obstacles. The work also evaluates continuous DTM-derived ruggedness metrics as complementary layers to characterise within-stand terrain variability and to assess their relationship to the obstacle-based, class-oriented framework. TLS is also used to support near-ground obstacle mapping via supervised point cloud semantic segmentation.

For steepness, the analyses quantify the sensitivity of slope and terrain complexity out-

comes to sensor type, grid resolution, and point-cloud processing choices. In this study, fine-scale terrain features and local steepness extremes were progressively smoothed as the grid size increased. Multi-sensor comparisons further show that slope-based category outcomes can vary markedly with resolution, and that apparent agreement at coarse grids can be misleading when spatial smoothing reduces the mapped extent and sometimes the occurrence of the steepest classes.

Overall, the dissertation concludes that remote-sensing-derived terrain metrics should be generated and interpreted with respect to specific planning objectives and operational constraints, including survey efficiency, processing demand, cost, and data volume. The results show that remote sensing can support a shift from stand- or polygon-level terrain descriptors toward spatially explicit raster- and point-cloud–based terrain layers that are more directly usable for routing, exclusion mapping, and operational planning. However, because terrain attributes are scale- and context-dependent, stand-level TCS thresholds require explicit translation before being applied to continuous surfaces (e.g., by defining aggregation windows, rescaling, or threshold adjustment). No single remote sensing platform is optimal across all applications. Systematic comparison instead supports evidence-based sensor and product selection that aligns analytical requirements with operational feasibility. Further research should validate these terrain layers against operational outcomes (e.g., machine performance, route choices, and soil disturbance) and develop decision-ready integration frameworks that translate multi-layer terrain information into actionable planning guidance.

Sommario

Remote Sensing–Based Terrain Analysis for Timber Harvesting Planning in Diverse Forest Environments

La sostenibilità della raccolta del legname con sistemi meccanizzati dipende da informazioni accurate sul terreno, poiché consentono di bilanciare efficienza operativa, convenienza economica, sicurezza delle macchine e tutela ambientale. I sistemi di classificazione del terreno (terrain classification systems, TCS) utilizzati in ambito forestale si basano tipicamente su descrizioni a scala di popolamento, rilievi di campo manuali e giudizio esperto. Sebbene efficaci per scopi di pianificazione generale, tali approcci offrono un dettaglio spaziale limitato, possono essere soggetti a bias dell'osservatore e risultano difficilmente scalabili su paesaggi forestali estesi.

Questa tesi indaga il potenziale dell'analisi del terreno basata su telerilevamento per quantificare attributi rilevanti per la pianificazione della raccolta meccanizzata e per valutarne l'affidabilità analitica e la fattibilità operativa in diversi contesti forestali, piattaforme sensoriali e scale spaziali. La ricerca è strutturata attorno a tre componenti fondamentali che costituiscono la base dei TCS consolidati: condizioni del suolo, rugosità superficiale e pendenza.

Attraverso casi studio basati su fotogrammetria Structure-from-Motion (SfM) da veicolo aereo senza pilota (UAV), laser scanning aerotrasportato (ALS), LiDAR da UAV e laser scanning terrestre e portatile (TLS), sono stati derivati indicatori del terreno e analizzate le loro relazioni con (a) la deformazione del suolo indotta dal passaggio delle macchine, (b) gli ostacoli prossimi al suolo e (c) la ripidità e la complessità del terreno, su diverse scale spaziali.

Per le condizioni del suolo, il lavoro ha esaminato se la prossimità ai ceppi e ai proxy di rinforzo ceppo/radici spieghi la variabilità intra-sito nello sviluppo delle ormaie in condizioni di traffico ripetuto, valutando inoltre modelli per la previsione della profondità delle ormaie. Modelli che combinano predittori derivati dal terreno con variabili operative hanno ottenuto valori di $R^2 = 0.69–0.85$ nella previsione della profondità osservata delle ormaie in diversi passaggi delle macchine e in configurazioni di tracciato.

Per la rugosità superficiale, la ricerca operazionalizza il concetto di rugosità basato sugli ostacoli, utilizzato nella classificazione forestale del terreno, combinando la fotogrammetria da UAV con la segmentazione degli ostacoli basata su deep learning. Nel contesto di studio ad alta visibilità, ciò ha consentito un tasso di rilevamento degli ostacoli del 95.6% (86/90) rispetto agli ostacoli marcati in campo. Il lavoro valuta inoltre indici continui di ru-

gosità/asperità derivati da modelli digitali del terreno (DTM) come livelli complementari per caratterizzare la variabilità intrapopolamento e analizzarne la relazione con l'impostazione basata su classi. Il TLS è inoltre impiegato per supportare la mappatura degli ostacoli prossimi al suolo mediante la segmentazione semantica supervisionata delle nuvole di punti.

Per la pendenza, le analisi quantificano la sensibilità degli esiti relativi a pendenza e complessità del terreno al tipo di sensore, alla risoluzione della griglia e alle scelte di elaborazione delle nuvole di punti. In questo studio, le caratteristiche del terreno a scala fine e gli estremi locali di ripidità risultano progressivamente smussati con l'aumentare della dimensione della griglia. I confronti multi-sensore mostrano inoltre che gli esiti di categorizzazione basati sulla pendenza possono variare sensibilmente con la risoluzione, e che un'apparente concordanza a griglie grossolane può essere fuorviante quando lo smoothing spaziale riduce l'estensione mappata—e talvolta anche l'occorrenza—delle classi più ripide.

Nel complesso, la tesi conclude che gli indicatori del terreno derivati dal telerilevamento devono essere generati e interpretati in relazione a specifici obiettivi di pianificazione e a vincoli operativi, tra cui l'efficienza di rilievo, la domanda di elaborazione, i costi e il volume dei dati. I risultati mostrano che il telerilevamento può supportare un passaggio da descrittori a scala di popolamento o poligonale verso layer del terreno spazialmente espliciti, basati su raster e nuvole di punti, più direttamente utilizzabili per l'instradamento, la mappatura di esclusione e la pianificazione operativa. Tuttavia, poiché gli attributi del terreno dipendono da scala e contesto, le soglie TCS definite a scala di popolamento richiedono una traduzione esplicita prima di essere applicate a superfici continue (ad es. definendo finestre di aggregazione, riscalando gli indicatori o adattando le soglie). Nessuna singola piattaforma di telerilevamento è ottimale per tutte le applicazioni; un confronto sistematico consente invece una selezione di sensori e prodotti basata su evidenze, allineando i requisiti analitici alla fattibilità operativa. Ulteriori ricerche dovrebbero validare questi layer rispetto agli esiti operativi (ad es. prestazioni delle macchine, scelte di percorso e disturbo del suolo) e sviluppare framework di integrazione "decision-ready" che traducano le informazioni multi-layer in indicazioni operative utilizzabili.

Opsomming

Remote Sensing–Based Terrain Analysis for Timber Harvesting Planning in Diverse Forest Environments

Volhoubare gemeganiseerde houtontginning is afhanklik van akkurate terreininligting om operasionele doeltreffendheid, kostedoeltreffendheid, masjienveiligheid en omgewingsbeskerming te balanseer. Konvensionele terreinklassifikasiestelsels (TKS) wat in bosbou gebruik word, is tipies gebaseer op standvlakbeskrywings, handmatige veldopnames en deskundige oordeel. Alhoewel hierdie benaderings doeltreffend is vir breë beplanningsdoeleindes, bied dit beperkte ruimtelike detail, is dit vatbaar vir waarnemervoorkeur/-vooroordeel en is dit moeilik om oor groot, heterogene boslandskappe te skaal.

Hierdie proefskrif ondersoek die potensiaal van afstandwaarneming-gebaseerde terreinanalisemetodes om terreineienskappe wat relevant is vir gemeganiseerde ontginningsbeplanning te kwantifiseer, en om die analitiese betroubaarheid en operasionele haalbaarheid oor verskillende bosomgewings, sensorplatforms en ruimtelike skale te evalueer. Die navorsing is gestruktureer rondom drie kernterreinkomponente wat gevestigde TKS-raamwerke onderlê, naamlik grondtoestande, oppervlakruwheid en helling.

Met behulp van ’n reeks gevallestudies gebaseer op onbemande lugvaartuig (UAV) Struktuur-van-Beweging (SfM)-fotogrammetrie, vliegtuiglaserskandering (ALS), UAV-ligopsporing en afstandbepaling (LiDAR) en draagbare en terrestriële laserskandering (TLS), is terreinmetrieke afgelei en ontleed in verband met (a) masjiengeïnduseerde grondvervorming, (b) naby-grond hindernisse en (c) steilheid en terrein-kompleksiteit oor verskillende ruimtelike skale.

Vir grondtoestande het die studie ondersoek of boomstomp-nabyheid en boomstomp-/wortelversterkingsproksies binne-terrein variasie in spoorvorming onder herhaalde masjienverkeer kan verklaar, en dit het voorspellingsmodelle vir spoordiepte geëvalueer. Modelle wat terrein-afgeleide voorspellers met operasionele veranderlikes kombineer, het R^2-waardes van 0.69–0.85 behaal vir die voorspelling van waargenome spoordiepte oor verskillende masjienherhalings en roetekonfigurasies.

Ten opsigte van oppervlakruwheid fokus die navorsing op die operasionalisering van die hindernisgebaseerde ruwheidskonsep wat in bosbouterreinklassifikasie gebruik word, deur UAV-fotogrammetrie met diepleergebaseerde hindernissegmentasie te kombineer. In die hoë-sigbaarheid-gevallestudie het dit ’n hindernisopsporingskoers van 95.6% (86/90) teenoor veldgemerkte hindernisse behaal. Die werk evalueer ook deur-

lopende, DTM-afgeleide ruigheid-ruggedness-metrieke as aanvullende lae om binne-stand-veranderlikheid te karakteriseer en om die verhouding tot die hindernisgebaseerde, klasgeoriënteerde raamwerk te beoordeel. TLS is verder gebruik om naby-grond-hinderniskaartering te ondersteun deur middel van beheerde (supervised) semantiese segmentasie van puntwolke.

Vir steilheid kwantifiseer die ontledings die sensitiwiteit van helling- en terrein-kompleksiteit-uitkomste vir sensortipe, roosterresolusie en puntwolkverwerkingskeuses. In hierdie studie is fynskaalterreinkenmerke en plaaslike steilheidsuiterstes progressief uitgevlak namate die rooster growwer geword het. Multisensorvergelykings toon verder dat hellinggebaseerde kategorie-uitkomste merkbaar met resolusie kan wissel, en dat ’n skynbare hoë ooreenstemming by growwer roosters misleidend kan wees wanneer ruimtelike uitvlakking die gekarteerde omvang, en soms selfs die voorkoms, van die steilste klasse verminder.

Oor al die terreinkomponente heen toon die resultate dat afstandwaarneming ruimtelik deurlopende terreinvoorstellings moontlik maak wat binne-stand-veranderlikheid vasvang wat dikwels nie deur konvensionele stand- of poligon-gebaseerde klassifikasie-benaderings opgelos word nie. Terselfdertyd wys die ontledings dat terreineienskappe inherent skaal- en konteksafhanklik is, en dat standvlak-TKS-drempels nie direk op deurlopende afstandwaarnemingsoppervlaktes toegepas kan word sonder eksplisiete vertaling nie (bv. deur aggregasievensters te definieer, metrieke te herskaal of drempels aan te pas).

Ten slotte kom die proefskrif tot die gevolgtrekking dat afstandwaarneming-gebaseerde terreinmetrieke gegenereer en geïnterpreteer moet word in verhouding tot spesifieke beplanningsdoelwitte en operasionele beperkings, insluitend opname-doeltreffendheid, verwerkingsvereistes, koste en datavolume. Die resultate toon dat afstandwaarneming ’n skuif kan ondersteun vanaf stand- of poligonvlak-terreindeskriptors na ruimtelik eksplisiete raster- en puntwolk-gebaseerde terreinlae wat meer direk bruikbaar is vir roetebeplanning, uitsluitingskartering en operasionele beplanning. Geen enkele afstandwaarnemingsplatform is optimaal vir alle toepassings nie; sistematiese vergelyking maak eerder bewysgebaseerde sensor- en produkkeuse moontlik wat analitiese vereistes met operasionele haalbaarheid belyn. Verdere navorsing behoort hierdie terreinlae teen operasionele uitkomste (bv. masjienprestasie, roetekeuses en grondversteuring) te valideer en besluitgereed (decision-ready) integrasieraamwerke te ontwikkel wat veelvuldige terreinlae in toepasbare beplanningsriglyne vertaal.

Table of Contents

List of Figures

List of Tables

Nomenclature

Functions and variables

Functions

$\mathbf{1}(\cdot)$	Indicator function: equals 1 if the condition holds, else 0.
CCC	Concordance correlation coefficient (used for slope agreement/accuracy in the feasibility trade-off figure).
CRITIC	Criterion importance through intercriteria correlation (objective weighting method).
CV	Coefficient of variation (dimensionless variability measure).
Moran's I	Global spatial autocorrelation statistic (used for clustering of canopy-/void patterns).
NMAD_z	Normalised median absolute deviation of elevation differences (robust vertical-accuracy metric).
OA	Overall agreement (%) between sensor and reference slope categories.
OE	Overestimation error (%): proportion of pixels where $s_i > r_i$.
PEI	Processing efficiency index: raw data size divided by processing time (e.g., GB min^{-1} or GB h^{-1}).
RMSE	Root mean square error (general definition).
RMSE_h	Horizontal (planimetric) RMSE of GNSS control/checkpoints (m).
RMSE_x	Horizontal RMSE component in the x direction when reported (m).
RMSE_z	Vertical RMSE of elevations (m).
Roughness	Local standard deviation of DTM elevation in a defined window (m).
Slope	DTM slope/gradient magnitude (reported in degrees and/or percent).
TCI(px)	Terrain complexity index at pixel px: weighted combination of normalised slope, roughness, and topographic wetness index (TWI).
$\mathrm{TRI}^{*}_{r,R}$	Detrended TRI: residual ruggedness after removing linear dependence on slope.
$\mathrm{TRI}_{r,R}$	Terrain ruggedness index at resolution r and neighbourhood radius R (m).
TWI	Topographic wetness index (unitless).
UE	Underestimation error (%): proportion of pixels where $s_i < r_i$.

VARIABLES

$\alpha_{r,R}$	Intercept of the linear model relating $\mathrm{TRI}_{r,R}$ to slope_r (used for detrending).
β_0	OLS intercept in the absolute-error regression model.
β_j	OLS partial regression coefficient for predictor j in the absolute-error regression model.
$\beta_{r,R}$	Slope coefficient of the linear model relating $\mathrm{TRI}_{r,R}$ to slope_r (used for detrending).
ΔZ	Vertical offset/difference in elevation (m), as reported for georeferencing/validation.
CR	Compression ratio: processed output size divided by raw data size (unitless).
CV_X	Coefficient of variation of variable X: $\mathrm{CV}_X = \sigma_X/\mu_X$.
RDR	Raw data rate: raw data volume divided by survey time (e.g., MB min^{-1}).
$\mathrm{TWI}_{\mathrm{norm}}(\mathrm{px})$	Normalised topographic wetness index (TWI) at pixel px (scaled 0–1 using robust min–max, 2nd–98th percentiles).
μ_X	Mean of variable X (used in CV).
ρ_g	Target ground-point density for UAV LiDAR thinning (pulses m^{-2}).
$\rho_{r,R}$	Pearson correlation between slope_r and $\mathrm{TRI}_{r,R}$.
σ_k	Standard deviation of metric k (CRITIC contrast-strength term).
σ_X	Standard deviation of variable X (used in CV).
ε_i	Residual (error term) at pixel i in regression models.
C_k	CRITIC contrast measure for metric k: $C_k = \sigma_k \sum_{i=1}^{m}(1 - r_{ik})$.
d_i	Pixel-wise difference between test and reference values (e.g., $d_i = z_{i,\mathrm{test}} - z_{i,\mathrm{ref}}$).
D_{raw}	Raw data size (volume) for a dataset (e.g., GB).
E_i	Vertical error at pixel i: $E_i = Z_{\mathrm{sensor},i} - Z_{\mathrm{reference},i}$ (m).
m	Number of metrics used in CRITIC weighting (e.g., slope, roughness, topographic wetness index (TWI)).
$m_{i,\mathrm{ref}}$	Metric value from the reference dataset at pixel i.
$m_{i,\mathrm{test}}$	Metric value from the evaluated (test) dataset at pixel i.
n	Number of pixels/observations used in a computation (e.g., RMSE, OA).
$P_{33}^{(r)}$	33rd percentile threshold of TCI at resolution r (used to define low/moderate/high classes).
$P_{67}^{(r)}$	67th percentile threshold of TCI at resolution r (used to define low/moderate/high classes).
R	Neighbourhood radius used for TRI variants (m).
r	Raster grid resolution / cell size (m).
r_i	Reference slope category at pixel i ($\in \{A, B, C\}$).

$R_{\text{norm}}(\text{px})$	Normalised roughness at pixel px (scaled 0–1 using robust min–max, 2nd–98th percentiles).
r_{ik}	Pearson correlation coefficient between metrics i and k (CRITIC).
s_i	Sensor-derived slope category at pixel i ($\in \{A, B, C\}$).
$S_{\text{norm}}(\text{px})$	Normalised slope at pixel px (scaled 0–1 using robust min–max, 2nd–98th percentiles).
T_{proc}	Processing time for a workflow/dataset (e.g., minutes or hours).
w	Moving-window size used for local terrain metrics (e.g., 5×5 m).
w_k	CRITIC-derived weight assigned to metric k.
w_r	Weight of roughness in the TCI formula.
w_s	Weight of slope in the TCI formula.
w_t	Weight of topographic wetness index (TWI) in the TCI formula.
$Z_{\text{reference},i}$	Reference DTM elevation at pixel i (m).
$Z_{\text{sensor},i}$	Sensor/test DTM elevation at pixel i (m).

ACRONYMS, ABBREVIATIONS AND DEFINITIONS

AGL	Above ground level
ALOS	Advanced Land Observing Satellite
ALS	Airborne laser scanning
AR	Area ratio
ASPRS	American Society for Photogrammetry and Remote Sensing
CCC	Concordance correlation coefficient
CI	Cumulative influence
CMOS	Complementary metal–oxide–semiconductor
CNN	Convolutional neural network
CPU	Central processing unit
CR	Compression ratio
CRITIC	Criterion importance through intercriteria correlation
CRS	Coordinate reference system
CSF	Cloth simulation filter
CTL	Cut-to-length
CV	Coefficient of variation
DEM	Digital elevation model
DJI	DJI (Da-Jiang Innovations)
DL	Deep learning
DSM	Digital surface model
DTM	Digital terrain model
DTW	Depth to water
EOSDIS	Earth Observing System Data and Information System
EU	European Union
FOV	Field of view
GCP	Ground control point
GDAL	Geospatial data abstraction library
GIS	Geographic information system
GNSS	Global navigation satellite system
GPR	Ground-penetrating radar
GPS	Global positioning system
GPU	Graphics processing unit
HLS	Handheld LiDAR scanning
ICP	Iterative closest point
IQR	Interquartile range
ISTVS	International Society for Terrain-Vehicle Systems
KDE	Kernel density estimation
LAS	LAS (LiDAR point-cloud file format)

LAZ LAZ (compressed LAS LiDAR format)
LiDAR Light detection and ranging
LWB Long wheelbase
MAE Mean absolute error
MAPE Mean absolute percentage error
ML Machine learning
MLS Mobile laser scanning
MRE Mean relative error
NASA National Aeronautics and Space Administration
NGP Nominal ground pressure
NMAD Normalised median absolute deviation
NTC National terrain classification
OA Overall agreement
OE Omission error (underestimation)
OLS Ordinary least squares
PALSAR Phased Array type L-band Synthetic Aperture Radar
PEI Processing efficiency index
RAM Random-access memory
RAR Root area ratio
RDR Raw data rate
RF Random forest
RFE Recursive feature elimination
RGB RGB (Red Green Blue)
RMS Root mean square
RMSE Root mean square error
RRV Root reinforcement value
RTK Real-time kinematic
SAM Segment anything model
SAR Synthetic aperture radar
SD Standard deviation
SfM Structure from motion
SLAM Simultaneous localisation and mapping
SSR Residual sum of squares
SST Total sum of squares
TCI Terrain complexity index
TCS Terrain classification system
TLS Terrestrial laser scanning
TRI Terrain ruggedness index
TWI Topographic wetness index
UAV Unmanned aerial vehicle

UE	User's error (overestimation)
UTM	Universal transverse mercator
ViT	Vision transformer
VRM	Vector ruggedness measure

DEFINITIONS

DETRENDED TERRAIN RUGGEDNESS INDEX (TRI*) A slope-adjusted ruggedness metric defined as the residual component of $\mathrm{TRI}_{r,R}$ after removing its linear dependence on slope_r (i.e., $\mathrm{TRI}^{*}_{r,R}$), intended to isolate ruggedness not explained by slope.

OPERATIONAL FEASIBILITY The practical ability to acquire, process, manage, and apply remote sensing–derived terrain data under real-world forestry constraints. In this work, it includes survey efficiency and logistics, data availability and coverage, processing complexity and computational demands, data volume and storage requirements, cost structure, and the time and expertise needed to deliver terrain data products suitable for analysis and decision-support. Operational feasibility is considered separately from analytical accuracy.

SLOPE-BASED TERRAIN CATEGORY A categorical classification (A, B, C) derived from NTC slope classes [1] using fixed slope thresholds (A: ≤30%, B: 30–50%, C: ≥50%), used to represent terrain steepness and its implications for mechanised forestry operations.

TERRAIN TRAFFICABILITY The ability of terrain to support and permit vehicle movement, i.e. whether vehicles can traverse an area without immobilisation, excessive sinkage/rutting, or loss of traction. In this dissertation, terrain trafficability is treated as a coupled terrain–vehicle property influenced by ground conditions (e.g. soil strength/support capacity and moisture state), surface roughness/obstacles, and slope, in line with ISTVS standards [2, 3].

TERRAIN COMPLEXITY INDEX (TCI) A composite, normalised index combining slope, surface roughness, and topographic wetness using CRITIC weighting, designed to represent relative terrain complexity at the pixel scale. In this study, TCI is stratified into three relative classes (low, moderate, high) using percentile-based thresholds and is intended for comparative and scale-sensitive analysis rather than as an absolute measure of terrain difficulty.

Chapter 1

Terrain characterisation in forest operations

1.1 The role of terrain analysis in mechanised timber harvesting planning

Timber harvesting is a global, high-volume industrial activity, whose scale and intensity are increasingly assessed in terms of operational efficiency and environmental footprint [4]. Global roundwood removals reached approximately 4.0 billion m^3 in 2022, with industrial roundwood production concentrated heavily in Europe and North America, accounting for approximately 55% of this global industrial volume [5]. This reflects a long-term trend of intensification, particularly in the European Union (EU), where roundwood production has increased by approximately 26% since 2000 [6].

In these industrialised forestry regions, commercial timber supply relies predominantly on fully mechanised harvesting systems [7, 8], typically full-mechanised ground-based cut-to-length (CTL) on gentle to moderate ground and cable yarding on steep terrain, usually involving the extraction of whole trees, with log processing carried out at the landing site [9, 10]. In CTL, a harvester fells, delimbs, and cross-cuts stems at the stump, after which a forwarder transports logs to the roadside landing. Because both machines make repeated passes within the stand, primarily along extraction trails, CTL concentrates traffic on a limited trail network. Cable-assisted (tethered/winch-assisted) configurations can extend the operating envelope of ground-based machines by improving traction and stability on steeper ground [11]. Where ground-based operation remains infeasible, cable yarding is the established alternative for slopes beyond the practical limits of conventional systems [12, 13].

While fully mechanised CTL systems dominate in parts of Northern and Central Europe and North America, the degree of mechanisation varies widely across global forestry regions, with many countries still relying on motor-manual and semi-mechanised systems [14]. In Mediterranean forestry contexts, harvesting often involves mixed systems, such as motor-manual felling combined with machine-based extraction (e.g., skidders or tractors), reflecting the later adoption and more limited spread of fully mechanised harvesting systems [15]. Although these vehicles are often lighter than purpose-built forest machines,

their environmental impact is not necessarily lower. Repeated traffic with wheeled tractors can still cause substantial soil compaction and rut formation, particularly under moist soil conditions, and the severity of disturbance is strongly influenced by terrain and soil factors such as soil moisture, ground morphology, and traffic intensity [16, 17].

While mechanisation has enabled substantial increases in labour productivity and harvesting capacity [18] and, by reducing reliance on motor-manual chainsaw felling, can minimise operator exposure to high-risk felling conditions [19], key limitations are no longer driven mainly by labour inputs (e.g. manual felling capacity and operator fatigue), but instead by how machines interact with terrain and soil conditions. In operational terms, this interaction is commonly framed as terrain trafficability, which refers to the ability of terrain to support and permit vehicle movement, encompassing factors such as soil bearing capacity, slope, and surface roughness, as defined in standards from the International Society for Terrain-Vehicle Systems (ISTVS) [2].

Modern forest machines are heavy, complex assets whose dynamic loading of the ground surface affects both operability and environmental outcomes [20]. In ground-based operations, slope gradient, surface roughness, and soil bearing capacity can therefore limit mobility and fuel efficiency and increase the likelihood of traction loss, vehicle instability, and soil deformation [21, 22, 23, 24, 25]. Where these limitations confine traffic to a limited and repeatedly used extraction network, cumulative machine passes can intensify impacts on soil physical properties, leading to compaction, rutting, and erosion [26, 27].

These terrain-related constraints also have economic and safety consequences. Harvesting on steep or rough ground typically requires more demanding operating conditions and specialised technology, and is often associated with reduced productivity [28] and higher unit costs [29]. Difficult terrain can also increase the severity of machine-related incidents [30] and elevate operator exposure to whole-body vibration during the operations [11, 31, 32]. Consequently, managers face an ongoing optimisation problem: maintaining economic performance by maximising productive machine hours and minimising extraction distance, since longer distances increase cycle time, fuel use, and unit costs. At the same time, they must meet regulatory (and often certification) requirements that limit soil disturbance (e.g., maximum allowable disturbed area and rutting thresholds) and protect site quality [33, 34]. For example, in Finland, rutting thresholds used in operational monitoring define a logging track as damaged if it contains more than 1 m of rutting deeper than 10 cm on mineral soils or 20 cm on peatlands [35]. These trade-offs make terrain information a prerequisite for effective planning. It is essential to identify where machines can operate efficiently and safely, and where soil disturbance risk is likely to exceed defined thresholds.

To optimise harvesting activities within these constraints, planning is typically organised within a hierarchical framework that distinguishes strategic, tactical, and operational levels [36]. This hierarchy reflects a practical balance: long-term strategic planning over

large areas relies on aggregation and simplified representations, whereas short-term operational execution requires fine-scale, spatially explicit information.

Terrain information, together with forest inventory, infrastructure, and operational constraint data, is consumed differently at each level while consistently serving as a feasibility filter. At the strategic level, terrain information supports broad operability zoning and system allocation. In practice, this includes deciding between ground-based and cable-based systems. At the tactical level, terrain information is used to translate these constraints into stand-level layouts, such as the placement of landings and the preliminary design of extraction networks and access. This means choosing a landing on stable ground near the road and sketching the main forwarder trails connecting the stand to that landing, while avoiding steep or wet areas. At the operational level, the focus shifts to short-term routing and disturbance avoidance, where decisions must be made at the scale of individual machine passes [36].

At this point, terrain data resolution becomes critical for assessing machine trafficability at operational scales. A single-value "steep stand" classification may be sufficient for strategic zoning, but it can miss short slope breaks, rock outcrops, stumps, boulders, and wet microsites that ultimately govern on-the-ground machine mobility and disturbance risk.

1.2 Terrain classification systems (TCS): historical context and framework

To bridge the gap between strategic zoning and operational reality, forestry developed standardised classification systems. The origins of these approaches, however, are not unique to forestry. The methodological foundations for general land classification were historically laid in land resource survey studies and military mobility/trafficability research [37], where approaches established the fundamental methodologies for describing terrain in repeatable mapping units during the 1940s–1960s [38, 39]. These systems prioritised consistency, repeatability, and applicability across large heterogeneous landscapes.

Building on these foundations, terrain classification systems emerged in forestry to standardise the assessment of ground conditions, match equipment to site constraints, and anticipate operational costs and environmental risks. Early forest engineering research in the Nordic countries developed systematic terrain assessment methods to address this need. Haarlaa [40] work within the Inter-Nordic Forest Terrain Classification Project demonstrated that a small set of physical parameters, such as ground conditions, surface roughness, and slope, could systematically predict machinery mobility, productivity, and environmental risk across sites. This approach became the template for subsequent national systems in Sweden, Canada, South Africa and New Zealand during the 1980s and 1990s [1, 41, 42, 43, 44], each adapting the three-factor structure to local soil, climate and operational contexts.

Despite these regional adaptations, national TCS implementations converge on the same three core terrain factors: ground conditions and/or ground bearing capacity, surface roughness and slope gradient, with difficulty classes defined on standardised scales. However, the specific thresholds and measurement methods vary substantially between systems, reflecting differences in soil types, climate, and equipment capabilities in each region (Fig. 1.1).

1.2.1 The three-pillar terrain classification framework

To organise these shared concepts, the three-pillar terrain classification framework formalises the standard three-factor structure and makes explicit how each factor constrains operations. Each pillar corresponds to a distinct operational constraint and is measured through regionally adapted methods.

- **Ground conditions** (bearing capacity) reflect soil strength under machine loads and are strongly moisture-dependent. The same soil can shift from ”good” trafficability when dry to ”poor” when saturated, with increasing water content generally reducing strength in fine-textured soils [1, 37, 45] (Fig. 1.1a). Systems operationalise this moisture sensitivity in different ways. The South African TCS explicitly codes ground condition for three moisture states (dry, moist, wet) using three digits (e.g., ”145”), with ratings derived from soil type and clay content and adjusted for moisture status [1]. In the New Zealand framework, bearing capacity is measured directly using a drop-hammer cone penetrometer and expressed as the number of *blows* required to reach specified penetration depths (commonly reported to ∼50 cm). Lower blow counts indicate weaker ground (poorer bearing capacity), while higher blow counts indicate firmer ground [45]. Field guidance notes that areas averaging fewer than about 15 blows to ∼50 cm warrant consideration of low-ground-pressure or cable extraction options, whereas where more than 15 blows are required, ground extraction is assumed feasible (subject to soil-moisture interpretation, especially on clay soils) [45]. The New Zealand protocol also specifies how the penetrometer readings are taken (multiple readings per plot; record penetration depth if 50 cm is not reached after 50 blows) to support repeatability [45].

  In contrast, Swedish and Finnish classifications rely primarily on mapped/field-interpreted soil-type and moisture-regime indicators rather than on routine in-situ strength testing. At the same time, North American approaches commonly infer ground strength from soil texture/moisture regime and site indicators, such as vegetation/soil associations [37, 40, 42, 43].

- **Surface roughness** describes micro-scale terrain irregularities, such as rocks, boulders, mounds, and depressions, that influence machine travel speed, stability, and fuel consumption [1, 11]. In the national terrain classification systems summarised

in Fig. 1.1b, roughness is expressed (where directly comparable) as obstacle density (obstacles ha^{-1}), but the class thresholds and descriptive labels differ between systems. For example, both Sweden and South Africa define the lowest roughness class as *isolated* terrain with $<$ 40 obstacles ha^{-1} and a second class spanning 40–400 obstacles ha^{-1} [1, 43]. Sweden then defines *moderate* roughness as 400–4000 obstacles ha^{-1} and an *abundant* class as $>$ 4000 obstacles ha^{-1} [43]. South Africa similarly uses 400–2000 obstacles ha^{-1} for a *moderately frequent* class and $>$ 2000 obstacles ha^{-1} for a *frequent* class [1]. New Zealand uses the same obstacle-density breakpoints as Sweden (i.e., $<$ 40, 40–400, 400–4000, and $>$ 4000 obstacles ha^{-1}), but applies different descriptive labels, with the 400–4000 range spanning *slightly uneven* through *uneven/rough* and $>$ 4000 classified as *very rough* [45]. The USA (Maine) system shown here is not directly comparable because roughness is defined as a matrix combining obstacle height and density rather than obstacle density alone [42].

The operational impact of surface roughness is direct and quantifiable: rough, obstacle-rich terrain reduces machine travel speed, increases fuel consumption and elevates operator workload, thereby constraining which machinery types can operate effectively [11]. Across all systems, the classification principle is consistent: terrain with high obstacle density and large obstacles poses the greatest impediment to machinery mobility and typically requires either slower operation or more capable equipment, regardless of regional context.

- **Slope gradient** governs machine traction, rollover risk, and the operational threshold for switching from ground-based to cable or aerial extraction systems [11]. All systems classify slope into discrete gradient classes (level, gentle, moderate, steep, very steep), which are tied to machinery limits and safety regulations. Yet, the numeric thresholds show the most significant variation between systems (Fig. 1.1c). Swedish [43] systems define the highest slope class (class 5) as $>$ 50%, with class 4 covering 33–50%. The South African system [1] similarly includes a $>$ 50% upper class. New Zealand [45] also uses $>$ 50% as the highest class, whereas the Maine (USA) scheme [42] uses $>$ 45% as the upper slope class. Additionally, slope shape, whether regular, undulating, terraced, concave, or convex, further modifies operational feasibility and safety, with convex slopes presenting elevated rollover risks and concave configurations potentially trapping machines or severely impeding lateral movement [1].

Despite this framework's clear operational structure, traditional TCS exhibit several practical limitations when implemented consistently across large landscapes.

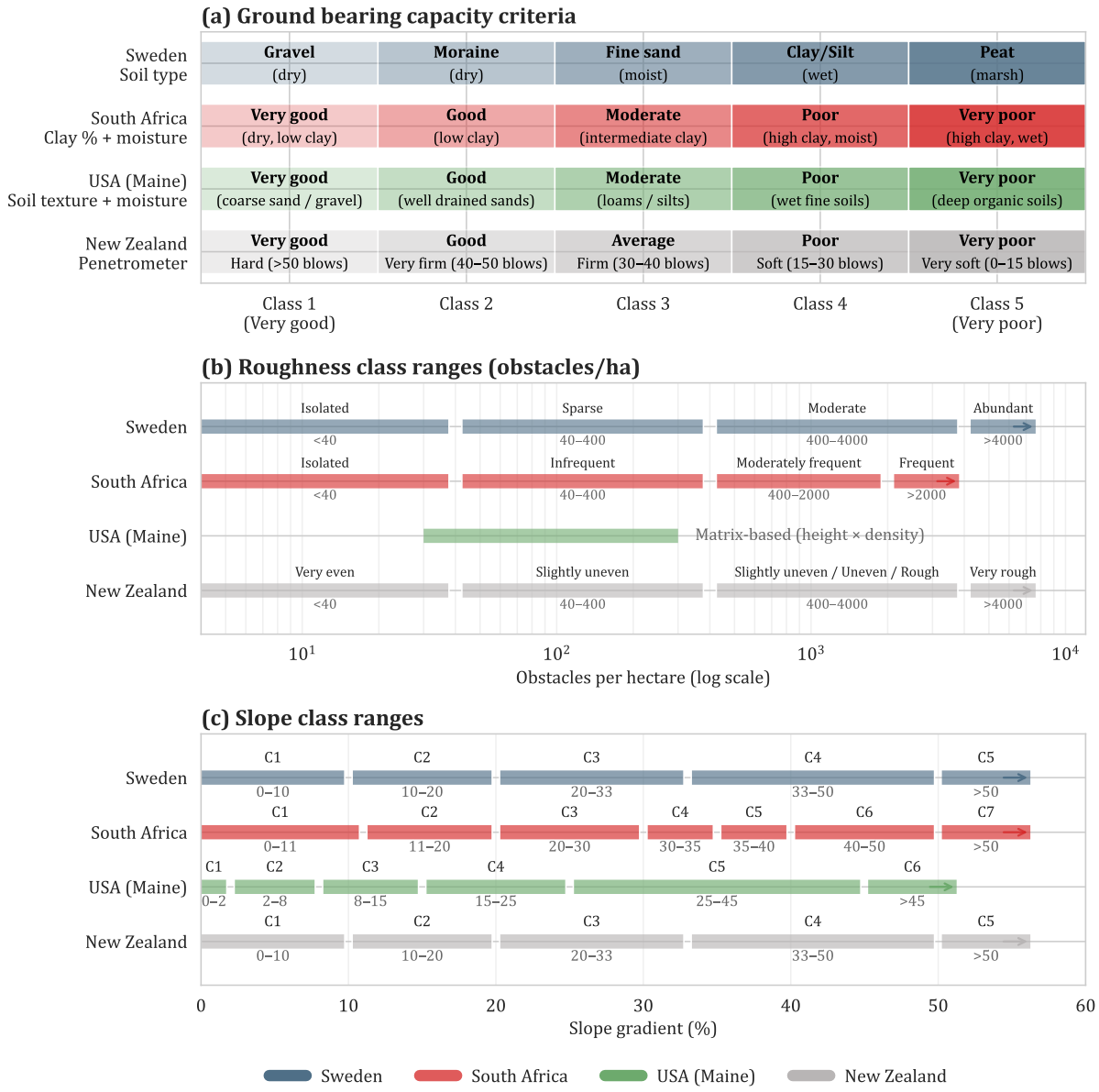


Figure 1.1: An overview of terrain classification criteria used in the four national frameworks. Panel (a) shows ground-condition classes describing soil strength and bearing capacity; (b) shows surface roughness classes expressed as obstacle density (obstacles per hectare; log scale); and (c) shows slope class ranges (slope gradient, %). For New Zealand, ground bearing capacity is based on cone penetrometer resistance reported as *blows to 50 cm*, i.e., the number of hammer impacts ("blows") required to drive the cone to a depth of 50 cm (higher blow counts indicate firmer ground). The systems are based on national terrain classification schemes for the USA (Maine) [42], South Africa [1], Sweden [43], and New Zealand [45].

1.2.2 LIMITATIONS OF TRADITIONAL TCS

Traditional TCS approaches rely heavily on manual field surveys, visual estimation, and expert judgement, which makes terrain mapping labour-intensive and difficult to scale across large landscapes [1, 45, 43]. Data collection typically requires in-field measurements for multiple terrain factors (e.g., bearing capacity via soil probing or penetrometer readings, roughness via obstacle counts, and slope along representative baselines) [1, 45]. As a re-

sult, terrain information is often limited to sampled transects or selected stands rather than continuous spatial coverage [1, 43, 45].

A second limitation is spatial representation. Terrain is commonly mapped as polygons intended to represent homogeneous units, often at scales of 100 m or larger, and assigned a single class for each factor [1, 43]. This discretisation imposes rigid boundaries that can mask gradual transitions and obscure operationally critical small-scale variability [1, 45]. Consequently, locally rough or weak ground may be averaged into a single class for an entire polygon, despite controlling machine mobility and disturbance risk in practice [1, 45].

Finally, reliability and temporal validity present persistent challenges. Visual ratings for roughness and slope are sensitive to observer judgement and mapping scale, requiring substantial training and calibration to achieve consistency among operators [1, 43]. In addition, trafficability can fluctuate with short-term changes in soil moisture and seasonal freezing, meaning that static terrain classes may only approximate conditions at the time of survey or mapping [1, 45]. This motivates terrain data sources that are more spatially exhaustive, less observer-dependent, and easier to update as conditions change [1, 45].

1.3 Remote sensing for terrain data acquisition

In recent decades, advances in remote sensing have transformed how forest terrain is mapped and quantified, supporting a wide range of operational and ecological applications [46, 47]. Unlike manual, unit-based mapping, remote-sensing–derived, spatially continuous data enable consistent topographic products across multiple spatial scales: from regional elevation models to fine-resolution datasets suitable for operational planning [48].

Platforms range from spaceborne (satellite) to airborne, unmanned aerial vehicle (UAV)-based, and terrestrial systems, each with trade-offs in coverage, achievable resolution, and performance under forest canopy [47, 49, 50, 51]. Across these platforms, sensors are commonly grouped into passive (which record reflected or emitted natural radiation) and active instruments (which emit their own signal and measure the return), and this distinction strongly influences ground-surface reconstruction in forest environments [52, 53].

1.3.1 Active sensors

Active sensors emit energy and measure the returned signal to capture three-dimensional geometry, most commonly as LiDAR (light detection and ranging) point clouds, which can be processed into digital terrain models (DTMs) and, where repeat acquisitions are available, elevation-change surfaces. In forest operations, LiDAR is widely used because it provides spatially continuous terrain and canopy information across scales, from regional baselines to stand-level planning layers [52, 47].

At landscape and planning scales, airborne laser scanning (ALS) provides area-wide LiDAR point clouds and derived terrain models (DTMs) that support access and extraction

planning and enable the derivation of susceptibility layers such as DTW (Depth-to-Water) [54] and TWI (topographic wetness index) [55] products [35, 47, 56]. These layers are often integrated into a geographic information system (GIS) and other workflows for tasks such as cost-surface routing and skid-trail network optimisation [57]. ALS-derived DTMs are also widely used for wet-area and trafficability mapping to delineate moisture-prone zones and support operational decisions [35, 47, 56]. In several EU countries, ALS also supports countrywide forest inventory and monitoring programmes, where area coverage is typically prioritised over very high point densities. Reported specifications are commonly around 1–4 points m^{-2}, depending on acquisition design [47, 58].

At finer operational scales, UAV LiDAR is often a more flexible option than ALS because it can be deployed on demand and repeated as needed to capture current site conditions. Compared with ALS, UAV LiDAR typically delivers denser point clouds over smaller areas, enabling a more detailed representation of microtopography for stand-level planning applications. It can capture small terrain features relevant to extraction planning, such as local obstacles and abrupt changes in the ground surface [59]. UAV LiDAR workflows have also been used to map post-harvest skid trails and soil disturbance using microtopographic signatures [60].

For detailed ground-surface reconstruction and disturbance quantification, terrestrial laser scanning (TLS) provides very dense, close-range point clouds. In forestry applications, TLS is often used as a high-accuracy reference for validating terrain products from airborne or UAV surveys and for quantifying fine-scale surface change in disturbance studies [61, 62]. For example, TLS has been used to map the spatial variability of soil disturbance along skid trails [63]. However, its main limitation is limited spatial coverage relative to airborne or UAV platforms.

Handheld and mobile LiDAR (HLS) enable efficient close-range mapping in forests, with SLAM (simultaneous localisation and mapping) commonly used to register scans when the global navigation satellite system (GNSS) is unreliable under canopy [64]. SLAM-based mobile scanning has been demonstrated on forest platforms to produce local 3D maps for extracting tree positions and stem attributes, while highlighting practical requirements such as loop closures and appropriate trajectory design to limit drift [64, 65].

Active sensing can also be integrated directly into operations through machine-mounted LiDAR, enabling rut-depth measurement during harvesting with centimetre-level agreement with manual reference measurements [66, 67]. Several studies further note that terrain products and wetness mapping are most informative when combined with dynamic predictors (e.g., hydrological modelling, weather history, and machine data) to address temporal variability in trafficability and rut development [24, 56, 68]. The main operational strengths and limitations of common active platforms are summarised in Table 1.1.

Table 1.1: Strengths and limitations of active sensing platforms for terrain acquisition in forest operations

Platform / sensor	Typical strengths for forest operations	Typical limitations / risks	Reference
ALS	Rapid, large-area coverage; consistent DTMs for strategic/tactical planning; enables DTW-style wet-area screening when high-res DTMs exist	Coarser representation of microtopography than UAV/close-range systems; DTM quality depends on ground returns and interpolation; lower-density products can miss/underestimate small details	[35, 47, 57]
UAV LiDAR scanning (ULS)	Very high local detail; flexible re-survey; strong for rugged terrain, obstacle detection, and stand-level terrain modelling	Smaller area per flight; workflow depends on GNSS/real-time kinematic (RTK) and/or ground control points (GCPs); high data volumes and computing/storage demands; higher acquisition and processing costs	[59]
TLS	Very dense point clouds; excellent for skid-trail microtopography and detailed disturbance mapping	Limited spatial extent; time-intensive field deployment; extremely dense point clouds leading to high processing and storage demands; high equipment and operational costs	[63, 69]
Handheld, mobile LiDAR (HLS, SLAM)	High point density at plot scale; efficient for detailed and rapid terrain mapping in small areas; useful where UAV flight is constrained	Drift/registration uncertainty under canopy; limited area coverage; higher per-hectare effort; large point clouds with substantial processing and storage demands; equipment and software can be costly	[70, 71, 72]
Machine-mounted LiDAR	Collects rut geometry during operations; supports routine monitoring without separate surveys	Mounting and viewing-geometry constraints; sensitivity to machine vibration and sensor placement; requires careful calibration, operational integration, and data quality checks	[66, 67]

1.3.2 Passive sensors

Passive sensors, on the other hand, record reflected radiation (e.g. RGB imagery) and infer surface geometry indirectly, most commonly through Structure-from-Motion (SfM) photogrammetry, making them effective for post-harvest disturbance mapping where canopy openness and ground visibility permit [52, 73]. UAV-based imagery has been widely used for post-harvest mapping of soil disturbance, including stand-level disturbance assessment and quantitative reconstruction of skid-trail deformation and rutting [70, 73, 74, 75].

In operational stand-level applications, orthomosaics and UAV-derived surface models (DSMs) have been used to map the spatial extent of disturbances and classify their severity across harvested areas [74]. Where surface geometry can be reliably reconstructed, UAV photogrammetry has also been used to quantify terrain deformation associated with extraction traffic, including estimating rut depth from SfM point clouds and derived elevation surfaces [73, 75]. In steep terrain, multi-temporal workflows combining post-harvest UAV surfaces with pre-harvest reference terrain enable direct quantification of soil displacement and cut–fill changes along machine trails, linking mapped disturbance patterns to measurable terrain change [70]. Across these studies, performance is consistently conditioned by canopy openness, understorey obstruction, and image geometry, which limit ground-surface visibility and therefore constrain where passive photogrammetric methods can provide complete and unbiased measurements [73].

At coarser scales, satellite-derived elevation products provide baseline terrain layers for regional context and broad mapping. However, their spatial resolution typically smooths the microtopography that governs machine–terrain interaction at the scale of skid trails, ruts, and local obstacles. This scale mismatch makes them most useful for strategic context. In contrast, operational mobility assessment and detailed soil-disturbance mapping generally require finer-resolution terrain products derived from UAV and LiDAR surveys [52, 47].

Nevertheless, satellite-based remote sensing (passive optical and active synthetic aperture radar, SAR) is widely used in forest operations for repeatable, large-area monitoring and inventory support. Multi-temporal observations can be used to update forest information systems with layers such as forest/non-forest, forest type, tree species distribution, and stand development stage. However, mapping performance depends strongly on the availability of imagery within appropriate phenological windows [76].

Open-access satellite data streams, such as Sentinel optical and SAR, and ALOS PALSAR (Phased Array type L-band Synthetic Aperture Radar) type SAR, have also been combined with machine-learning workflows to map forest structure variables (e.g., height, basal area, biomass proxies) at management-relevant resolutions, with demonstrated benefits from fusing optical indices, textures, and multi-frequency SAR features [77]. In operations oversight, satellites are commonly used for change detection and compliance monitoring, for example, to identify harvesting activity and delineate affected areas, leveraging repeat revisit and synoptic coverage [78]. For post-harvest soil-impact screening at broader scales,

coarse-to-moderate-resolution satellite imagery, such as Sentinel-2, has been evaluated alongside other readily available image sources for detecting extraction-related track patterns. However, reliability remains site- and visibility-dependent compared with higher-resolution UAV products [50].

1.3.3 Terrain metrics for forest operations

Regardless of the sensor or platform, remote sensing data used for terrain analysis are typically processed into gridded elevation surfaces, such as DSMs and DTMs, and, when image-based workflows are used, orthomosaics provide additional visual interpretation and feature mapping.

From these products, raster layers and terrain metrics are typically derived and used across application domains (e.g., geomorphology, where DTMs and related indices support landslide modelling and terrain-stability assessment) [79]. In forest operations, they are most commonly used to quantify deformation linked to soil disturbance, characterise surface roughness or ruggedness using continuous indices, and represent steepness and broader terrain complexity for planning and trafficability [52, 47].

1.3.3.1 Soil deformation and rutting

Ruts develop when wheel or track loads (and associated shear stresses under traction, braking, and turning) exceed the soil's bearing capacity, leading to plastic deformation and displacement of the upper horizons. Because rutting can impair machine mobility and access, alter surface hydrology and connectivity, and increase erosion and sediment-transport risk, it is widely used both as an environmental-impact indicator and as a practical quality-control variable in mechanised harvesting [47].

In Nordic practice, this operational relevance is reflected in monitoring frameworks that treat rut-depth thresholds as triggers for operational adjustments, such as re-routing traffic, adding slash, or postponing forwarding, and in decision-support efforts that aim to identify wet or weak areas before machines enter them [47, 80]. From a terrain-analysis perspective, rut-risk mapping is often framed as a soil-moisture-susceptibility problem, in which topographic indices approximate zones of water accumulation and shallow groundwater.

Two widely used metrics are the TWI and depth-to-water (DTW). TWI is a steady-state wetness proxy based on upslope contributing area and local slope [55]. DTW estimates the vertical distance from the terrain surface to the nearest mapped surface-water feature along a flow path, with low DTW values indicating locations more likely to be wet and weak [54]. In trafficability mapping, DTW is commonly thresholded to delineate potentially sensitive soils (e.g. using DTW ≤ 1 as a nominal "wet/sensitive" class) [47]. However, empirical evaluation shows that DTW performance can be site-dependent and should not be treated as a stand-alone predictor. In a survey of 16 Swedish clear-cuts, severe rutting (ruts deeper than 10 cm) was significantly associated with soil bearing-capacity class and traffic

intensity. In contrast, cartographic DTW alone was not significant in the pooled analysis. Nevertheless, DTW added value when interpreted jointly with soil information for road extraction and planning [80].

Additionally, an increasing portion of the work has evaluated remote sensing as a means of objectively measuring rut depth and mapping rutting extent and severity. UAV-based photogrammetry and close-range 3D reconstruction can capture rut geometry with centimetre-level agreement relative to field measurements under favourable visibility conditions [70, 73, 74, 81]. For example, close-range photogrammetry reproduced rut depth with a root-mean-square error (RMSE) of a few centimetres (2.07–3.84 cm across profiles). It supported volumetric disturbance estimates along short trail segments [70]. UAV photogrammetry similarly enables continuous spatial profiling of ruts. However, accuracy depends on acquisition design and rut morphology. A drone-based study reported high agreement with manual measurements (R^2 of 0.77–0.83) and noted underestimation at higher flight altitudes, as well as practical failure modes such as water-filled ruts [82]. In a skidding-operation case study, UAV-based rut estimation was reported with a mean absolute percentage error (MAPE) of 16.86% [75]. Beyond post-harvest UAV surveys, machine-mounted LiDAR has been tested as an in-operation monitoring option, producing rut-depth estimates with RMSE < 3.5 cm when validated against manual reference measurements [66]. Together, these approaches provide spatially explicit deformation metrics that can support post-harvest assessment and, notably, help validate terrain-based trafficability indicators used in planning [47, 74].

1.3.3.2 Surface roughness and terrain ruggedness

Terrain roughness in forestry can be represented in two conceptually different ways. In traditional TCS, roughness is commonly described as an obstacle-based property, characterised by the frequency and height of discrete impediments such as rocks and boulders [1, 43]. In contrast, DEM-based indices such as the terrain ruggedness index (TRI), vector ruggedness measure (VRM), and area ratio (AR) represent continuous measures of surface variability computed from elevation neighbourhoods, rather than explicit obstacle counts. Because these approaches quantify different aspects of the ground surface, they are not directly interchangeable, and DEM-based ruggedness indices are sensitive to analysis scale (neighbourhood size and DEM resolution) and to slope-related effects [83, 79].

Given this conceptual mismatch, recent work has also begun to bridge these concepts by using high-resolution UAV terrain surfaces to automate obstacle-focused roughness mapping that aligns with the logic of traditional, obstacle-based roughness classes [83]. In parallel, other studies treat ruggedness indices as general-purpose descriptors of surface complexity and show that their outputs are strongly shaped by methodological choices, including the input elevation model, raster resolution, and neighbourhood definition [79].

However, recent methodological work shows that widely used DEM-based metrics, especially TRI, can be strongly influenced by local slope (and may behave like a slope proxy

on smooth but inclined surfaces) unless slope/trend effects are explicitly addressed [84]. Accordingly, index choice and parameterisation should be matched to the operational question and mapping scale.

1.3.3.3 Steepness and terrain complexity

Steepness is a fundamental factor in forest operations because it affects traction, stability, and the practical applicability of different extraction systems. In steep terrain, conventional ground-based systems are often constrained to lower slopes, while cable-based systems become the dominant option as gradients increase. For example, steep-slope operations in mountainous regions are strongly associated with cable yarding, and commonly cited safe operating ranges for standard ground-based systems are typically below about 30–40% slope [12]. Empirical studies of cable extraction in southern Italy illustrate the steep context in which cable systems are applied, with reported average slopes frequently exceeding 50% and ranging from 59% to 78% across study sites [13].

In terrain modelling, steepness is typically represented as slope derived from DTMs. It is often combined with complementary descriptors such as curvature, local relief, and wetness/flow-related indices in GIS-based suitability and planning workflows. The usefulness of slope products is scale-dependent: higher-resolution DTMs better capture short, steep segments and slope breaks that can control route feasibility, whereas coarser DEMs smooth these transitions, potentially misrepresenting critical sections for machine routing and constraint mapping [85]. In operational decision support, slope layers are therefore used together with other mapped constraints, particularly wet-area/trafficability proxies such as DTW/TWI-type indices, because steepness interacts with soil moisture and bearing capacity in shaping machine performance and rutting risk [47, 56]. Recent work further emphasises the importance of combining terrain morphology with time-varying drivers (e.g., weather history and modelled soil moisture) and operational data streams to improve the explanation of performance variation in mechanised CTL systems [86].

Collectively, these terrain metrics describe how the ground surface responds to machine loading (deformation and rutting), how uneven or rugged it is at operational scales, and how steepness and terrain complexity constrain the choice and performance of harvesting systems [52, 47]. Yet, their reliability and operational usefulness depend strongly on sensor type, ground visibility, and analysis scale, as well as on how well DEM-based indices capture the constructs embedded in traditional terrain classification systems [79, 84, 83]. These dependencies motivate a structured evaluation of remote sensing-based terrain analysis across the three terrain pillars, together with an assessment of the practical trade-offs (survey effort, processing demands, and data quality) that determine whether these methods are feasible in forest operations.

1.4 Research aims and objectives

1.4.1 Overall research aim

The primary aim of this research was to evaluate the potential of remote sensing–based terrain analysis to objectively quantify terrain attributes relevant to timber harvesting planning, and to assess their reliability across different forest environments, sensors, and spatial scales.

The three-pillar TCS framework provides the conceptual structure of the thesis, where each pillar is addressed through a dedicated study in a distinct forest environment (Fig. 1.2).

1.4.2 Specific research objectives

RO1 *(soil conditions)*: To quantify terrain attributes governing machine–terrain interaction, with particular emphasis on ground conditions and surface deformation, using high-resolution remote sensing and field-referenced data.

RO2 *(surface roughness)*: To develop and evaluate remote sensing–based methods for deriving surface roughness and ground-level obstacle information consistent with terrain classification system concepts, and to generate spatially explicit roughness products for use in mechanised harvesting planning.

RO3 *(steepness and terrain complexity)*: To assess how sensor type, spatial resolution, and terrain complexity influence DTM-derived terrain representation and slope metric accuracy, and to evaluate how these effects propagate into slope-based terrain categorisation across sensors and spatial scales.

RO4 *(operational feasibility)*: To evaluate the operational feasibility of different remote sensing platforms and terrain data products in forestry applications, considering survey efficiency, processing demands, data volume, cost structure, and logistical constraints.

1.4.3 Research questions

The following research questions operationalise these objectives and guide the thesis structure.

RQ1 *(soil conditions)*: To what extent can remote sensing–based terrain analysis quantify ground conditions and surface deformation, and support modelling of machine–soil interaction responses in forest operations?

RQ2 *(surface roughness)*: To what extent can remote sensing–based terrain analysis derive operationally meaningful surface roughness and obstacle metrics that are consistent with established TCS concepts under challenging site conditions?

RQ3 *(steepness and terrain complexity):* How do sensor type, spatial resolution, and terrain complexity affect DTM-derived terrain representation, slope estimation, and the consistency of slope-based terrain categorisation?

RQ4 *(operational feasibility)*: What practical constraints and trade-off affect the operational feasibility of different remote sensing platforms and terrain data products?

1.4.4 Research scope and limitations

This dissertation focuses on the application of active and passive remote sensing technologies for high-resolution terrain characterisation in the context of forest operations. The research is based on case studies from three distinct forest environments: boreal peatlands in Finland, a post-fire commercial pine plantation in South Africa, and mixed mountain forest in Italy.

The scope of the research is delimited as follows:

- **Forest machinery:** The analysis focuses on terrain interactions (trafficability, obstacle traversability, and soil deformation) specific to ground-based CTL machinery. While the resulting high-resolution terrain models could also inform other harvesting systems, the mobility constraints analysed here are specific to ground-based wheeled and tracked machinery.
- **Remote sensing sensors:** The research primarily evaluates UAV-based photogrammetry (SfM) and LiDAR-based (ALS, UAV LiDAR, TLS and HLS) platforms. A NASA satellite-derived DEM is included only in a limited comparative context. Beyond this baseline, other satellite-based remote sensing products and SAR approaches are not analysed.
- **Terrain attributes:** Terrain analysis focuses on the three core components commonly used in TCS: soil conditions (expressed through surface deformation and bearing-capacity proxies), surface roughness (ground-level obstacles), and slope gradient (examined primarily from a methodological perspective rather than direct trafficability modelling).

The study has several limitations that should be considered when interpreting the results:

- **Integrated decision-support and performance modelling:** The research evaluates terrain attributes and terrain classifications as planning-relevant inputs. It does not, however, develop an integrated operability model or optimise harvesting productivity, fuel/energy consumption, or machine routing and extraction layout.

- **Operational integration:** The analysis is based on survey-grade datasets used for planning and post-processing. Real-time, machine-mounted sensing (e.g., onboard cameras or SLAM-based systems) and integration with machine control data (e.g., controller area network (CAN) bus) for in-operation terrain assessment are not evaluated.
- **Economic evaluation:** Operational feasibility is assessed in terms of survey effort, data volume, and processing requirements. A formal economic or financial cost–benefit or return-on-investment analysis, including detailed expenses, labour time, transport, and field logistics, is therefore outside the scope of this study. .
- **Soil properties and proxies:** Direct, systematic laboratory analysis of soil physical, chemical, or biological properties is not undertaken, and no comprehensive geotechnical testing campaign is implemented. Ground conditions are instead represented through surface deformation and terrain-based proxies (and, where available in individual case studies, in situ strength indicators such as shear vane measurements) rather than complete geotechnical soil-strength characterisation.
- **Temporal limitations:** Analyses represent conditions at the time of data acquisition and were conducted primarily under snow-free conditions. Seasonal effects (e.g., frozen soils, snow cover) and time-varying soil moisture and trafficability dynamics are not modelled explicitly.
- **Model generalisability:** Machine learning workflows were developed and validated using site-specific training data. Therefore, the transfer to other forest types, terrain conditions, or sensor configurations may require additional training or recalibration.

1.5 Overall workflow

Building on the three-pillar structure outlined above, Fig. 1.2 summarises how the individual studies connect through a common methodological workflow, from data acquisition and terrain product generation to cross-sensor evaluation and operational feasibility assessment.

1.6 Thesis outline

The thesis is organised around the three core terrain classification pillars (ground conditions, surface roughness, and slope/steepness), with each analytical chapter addressing one pillar (or its extension) and contributing to an integrated workflow and feasibility evaluation.

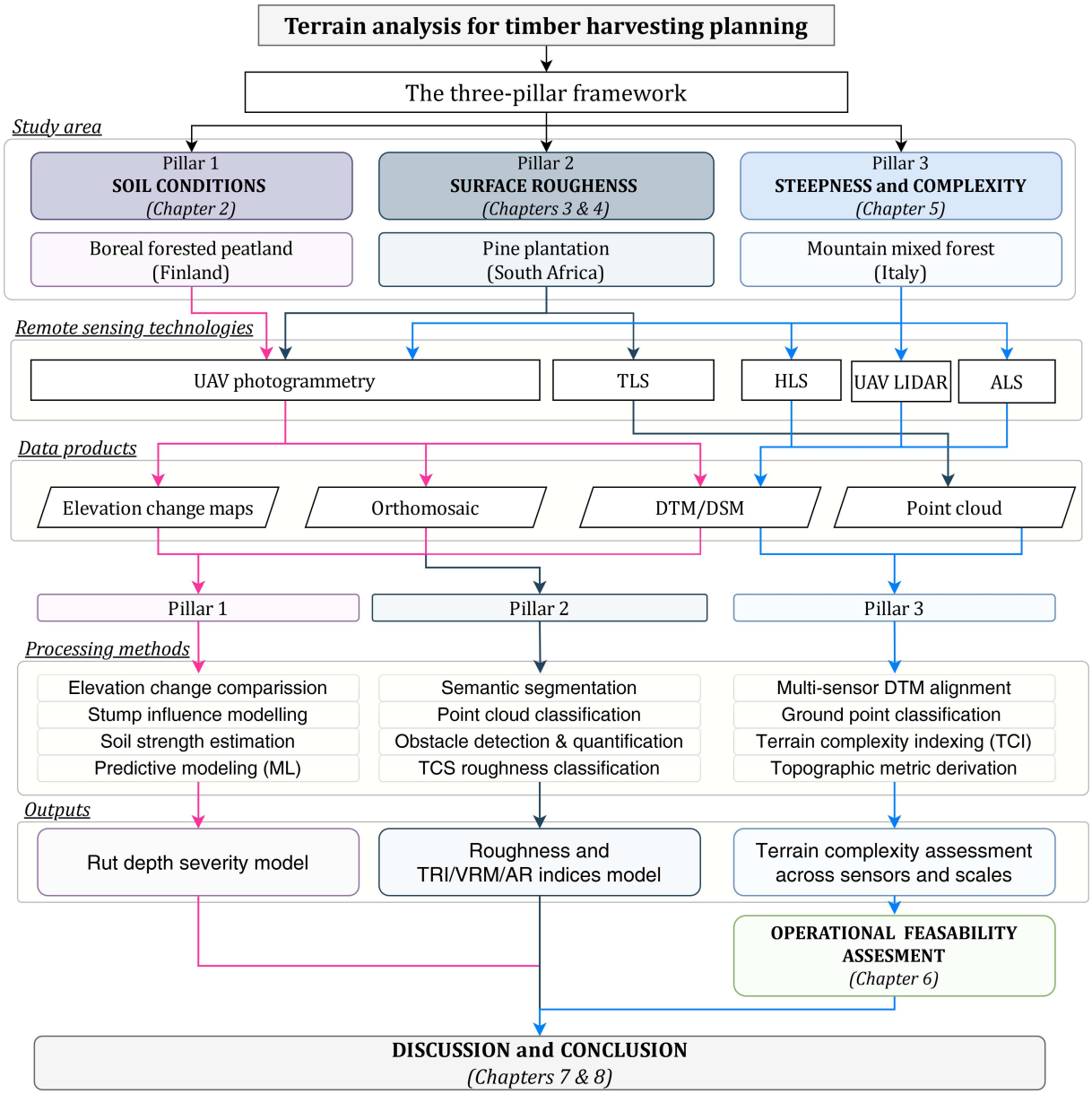


Figure 1.2: An overview of the overall research workflow connecting remote sensing technologies, terrain classification pillars, processing methods, and operational feasibility assessment.

- **Chapter 1** introduces the research context, the terrain classification framework, the role of remote sensing in terrain data acquisition, and the study's aims, scope, and structure.
- **Chapter 2** investigates the first pillar, ground conditions, by modelling machine-induced soil deformation and the mitigating effect of root reinforcement in peatland forests.
- **Chapter 3** addresses the second pillar, surface roughness, by developing a UAV photogrammetry workflow to detect and classify individual obstacles in post-fire terrain.
- **Chapter 4** extends the surface-roughness analysis by evaluating high-density TLS data to detect obstacles under canopy cover (preliminary results).

- **Chapter 5** focuses on the third pillar, steepness and terrain complexity, by comparing multi-sensor data (ALS, UAV LiDAR, HLS, SfM) to assess accuracy and classification reliability in steep mountain terrain.
- **Chapter 6** synthesises the operational feasibility of these methods, evaluating the trade-off between data quality and volume, survey effort, and processing requirements.
- **Chapter 7** discusses the broader implications of the findings for forest operations and future research directions.
- **Chapter 8** provides the final conclusions and summarises the research contributions.

Chapter 2

Modelling machine-induced soil deformation in forest soils using stump proximity and machine learning

2.1 Abstract

Soil deformation is a key challenge in sustainable timber harvesting, particularly in low-bearing-capacity environments. In mechanised forestry, this issue is particularly pronounced in peatlands, where rutting results from displacement and root shearing within the soft, organic substrate. Although tree roots are known to reinforce soil, the role of stump-root systems in mitigating rut formation remains underexplored. This study investigates the influence of stump presence on rut depth using UAV-based digital terrain models (DTMs), manual field measurements, spatial modelling, and machine learning techniques. Unmanned Aerial Vehicle (UAV) rut depth measurements were first compared with manual data, showing slightly lower values in deeper ruts, particularly in curved trails, with average discrepancies of around 3 cm. Statistical analysis confirmed that cumulative stump influence significantly reduced rut depth, with a small to medium effect in straight trails (ε^2 = 0.04–0.20) and a moderate to significant impact in curved trails (ε^2 = 0.02–0.32). Machine learning models achieved high predictive accuracy (R^2 = 0.69–0.85), identifying stump-related variables and soil shear modulus as key predictors of rut formation. These findings underscore the importance of incorporating stump-root reinforcement into forest planning to optimise machine path selection and minimise soil disturbance. Future research should refine species-specific reinforcement models and explore advanced root mapping techniques, such as ground-penetrating radar (GPR), to enhance decision-support tools for sustainable forestry.

2.2 Introduction

Forests are a key source of renewable resources, essential for timber production and economic sustainability [87, 88]. The increasing demand for sustainably sourced timber has

intensified efforts to optimise wood mobilisation while ensuring long-term ecological balance [89, 90]. In many parts of Northern Europe, the Baltic states, and Russia, a substantial share of timber production takes place on organically rich soils, including peatlands, which are widespread across these regions and characterised by high organic matter content [91].

These ecosystems are known for their high carbon storage capacity and waterlogged conditions [92], which pose unique challenges for timber production. Historically, peatlands have often been managed through drainage to increase forest productivity and timber supply, thereby supporting tree growth and facilitating easier harvesting processes [93]. It is estimated that approximately 15 million hectares (Mha) of peatlands have been drained for forestry purposes in the temperate and boreal regions, with about 10 Mha in the Baltic Sea Region (Finland, Sweden, Estonia, Latvia, Lithuania, and parts of Russia) alone [91, 94]. For instance, in Finland, peatlands cover approximately 8.8 Mha, accounting for 33% of forestry land, with slightly more than half (4.6 Mha) being drained for forestry. These drained peatlands contribute significantly to timber production, accounting for 25% of the total annual stem volume increment of 104 Mm^3, with a mean annual increment of 4.6 m^3 ha^{-1} [95]. The share of managed peatland forests is highest in southern and eastern Finland, where 88-90% of the peatland area has been drained, compared to just to 23% in Lapland [93].

Despite their economic importance, drained peatlands pose significant challenges for forestry operations due to their distinctive physical and ecological characteristics, including low bulk density and high porosity [96]. Unlike mineral soils, peatlands are composed primarily of accumulated organic material, formed over thousands of years under water-saturated, oxygen-poor conditions [94]. This results in a multi-layered structure, typically consisting of a surface layer (*acrotelm*) that is up to several decimeters thick. This layer is porous and highly permeable to water, containing living mosses, roots, and decomposing organic material. The deeper or sub-surface layer (*catotelm*) consists of variably decomposed, waterlogged peat with very low shear strength [94, 97, 98].

Peat deposits vary in thickness, with boreal forested peatlands typically having an organic layer exceeding 30 cm [94, 98, 99]. As the peat depth increases, its bearing capacity decreases, making the soil highly susceptible to deformation under external loads [22, 100]. When heavy forestry machinery operates on these soft, saturated soils, the high ground pressure, combined with low structural integrity, leads to soil displacement, compaction, and rut formation [101]. These disturbances are particularly severe under poor trafficability conditions, where repeated machine passes may further destabilise the peat [26, 49, 102], leading to long-term soil degradation, hydrological disruption, and reduced tree growth.

Research has shown that rut formation is strongly influenced by soil strength, traffic frequency, and trail design, with deeper ruts forming as shear strength decreases and the number of machine passes increases [103, 104, 105]. Forwarders equipped with wider

trails or additional wheels help mitigate rut depth by distributing ground pressure more effectively. For instance, trials with 700 mm wide tyres showed a significant reduction in rut depth compared to 600 mm configurations under equivalent loads [106].

However, while they reduce rutting during the early machine passes, rut formation continues to progress with repeated traffic, particularly on soft soils, where continued compaction and displacement occur [107, 108]. Trail design also plays a role, as wider trails (6 m) allow operators to avoid deepening existing ruts, though further widening beyond this offers little additional benefit [103]. Furthermore, modern forwarders are more favourable in minimising rut formation compared to lighter forwarders from the 1980s, likely due to the use of bogie axles in the front and improved trail design, which helps distribute weight more effectively and reduce soil pressure [101].

In recent years, to enhance insight into soil disturbance caused by forestry machinery, numerous studies have explored the application of remote sensing techniques, particularly UAV-based photogrammetry and machine learning. Drone-based measurements provide highly accurate rut depth estimates that closely match manual measurements while enabling large-scale, efficient assessments [82]. Deep learning (DL) applied to UAV imagery achieved 79.5% accuracy in detecting ruts, with the highest precision for severe rutting [109]. Additionally, Depth-to-Water (DTW) maps help predict rut severity, with 71% of severe ruts occurring in areas with DTW values below 1 m [23].

While these advances in remote sensing improve the ability to monitor rut formation, they do not address the underlying factors that influence soil resistance to deformation. One critical but often overlooked aspect is the role of root systems in reinforcing soil mechanical strength. Roots improve soil stability through tensile resistance, root-soil interlocking, and mechanical support, which together increase shear resistance, distribute stress, and bind soil particles [110, 111, 112]. These factors directly influence the soil's shear modulus, which quantifies its resistance to deformation, which determines its structural integrity and ability to resist rut formation [113]. Because roots contribute to shear strength, they also play a fundamental role in soil reinforcement, with their effectiveness depending on factors such as root density, root distribution, and tensile strength [114].

These characteristics differ among tree species and soil types, resulting in variations in root reinforcement values [115]. In boreal forests of Finland, Sweden, and the Baltic states, the dominant species, Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*), both contribute to soil reinforcement, but in different ways. Scots pine typically develops a deep root system with strong tap roots, improving stability in mineral soils. However, on wet peatlands, it may also adopt a shallower root structure. Norway spruce, on the other hand, has a shallower, plate-like root system that reinforces the upper peat layer [115, 116].

Root reinforcement is typically evaluated using empirical or mechanistic models that assess the contribution of roots to soil stability [114]. Most studies in recent decades have followed the approach of [117] and [118], which treat root bundles as adding cohesion to the soil. These models require parameters such as root tensile strength, root area ratio

(RAR), and root distribution, which are obtained from field and laboratory tests [114, 119]. At the same time, advanced models such as the root mundle model with Weibull survival (RBMw) [112] and RootMap [114] improve predictions by incorporating progressive root failure. They primarily focus on slope stability and do not account for the influence of stump roots on soil response to machine traffic.

Some studies have attempted to quantify the spatial effects of stump-root systems [120] and have found that stumps reduce variability in rut depth along skidding trails. However, comprehensive spatial modelling approaches remain underdeveloped, and no existing model fully integrates stump-root reinforcement, soil properties, and machine loading into a predictive model. There is a clear need for predictive tools that support machine path planning, enabling managers and operators to avoid high-risk areas and minimise soil disturbance. Improved management strategies are also essential for optimising forestry operations on sensitive soils.

This study aims to address that gap by developing a spatially explicit model that integrates stump locations and characteristics, soil and environmental parameters, and machine-induced depth values. Specifically, this study aims to (a) evaluate the accuracy of UAV-based digital terrain models (DTM) by comparing them with manual rut depth measurements and (b) develop a spatial decay model that interpolates stump diameter and root reinforcement effects over distance. Additionally, the study utilised a machine-learning model to assess key predictors of rut depth and to evaluate how stump-related variables influence soil deformation in straight and curved trail configurations. By further integrating this approach with DTW maps, the model could improve decision-making in drained peatlands, reducing the environmental impact of forestry operations.

2.3 MATERIALS AND METHODS

This chapter quantifies machine-induced rutting on peatland soils and evaluates the extent to which stump–root systems reduce soil deformation. Methods combine UAV photogrammetry (multi-pass DTMs), manual rut-depth measurements, stump mapping, spatial stump-influence modelling, and Random Forest prediction. The workflow includes two linked components: (1) validation of UAV-derived rut depth against manual measurements and (2) rut-depth prediction from soil, machine, and stump-derived predictors (Fig. 2.1).

2.3.1 STUDY AREA

The research was conducted in a peatland area near Tiilikkala in the municipality of Rautavaara, Northern Savonia, Finland (61° 32’ 0” N, 26° 50’ 0” E, UTM zone 35N) (Fig. 2.2). The Rautavaara region is generally flat, with elevations ranging from sea level to about 100-200 meters above sea level. The landscape is primarily natural, with vast peatlands, extensive forests, and numerous lakes, including the nearby Lake Pielinen [97].

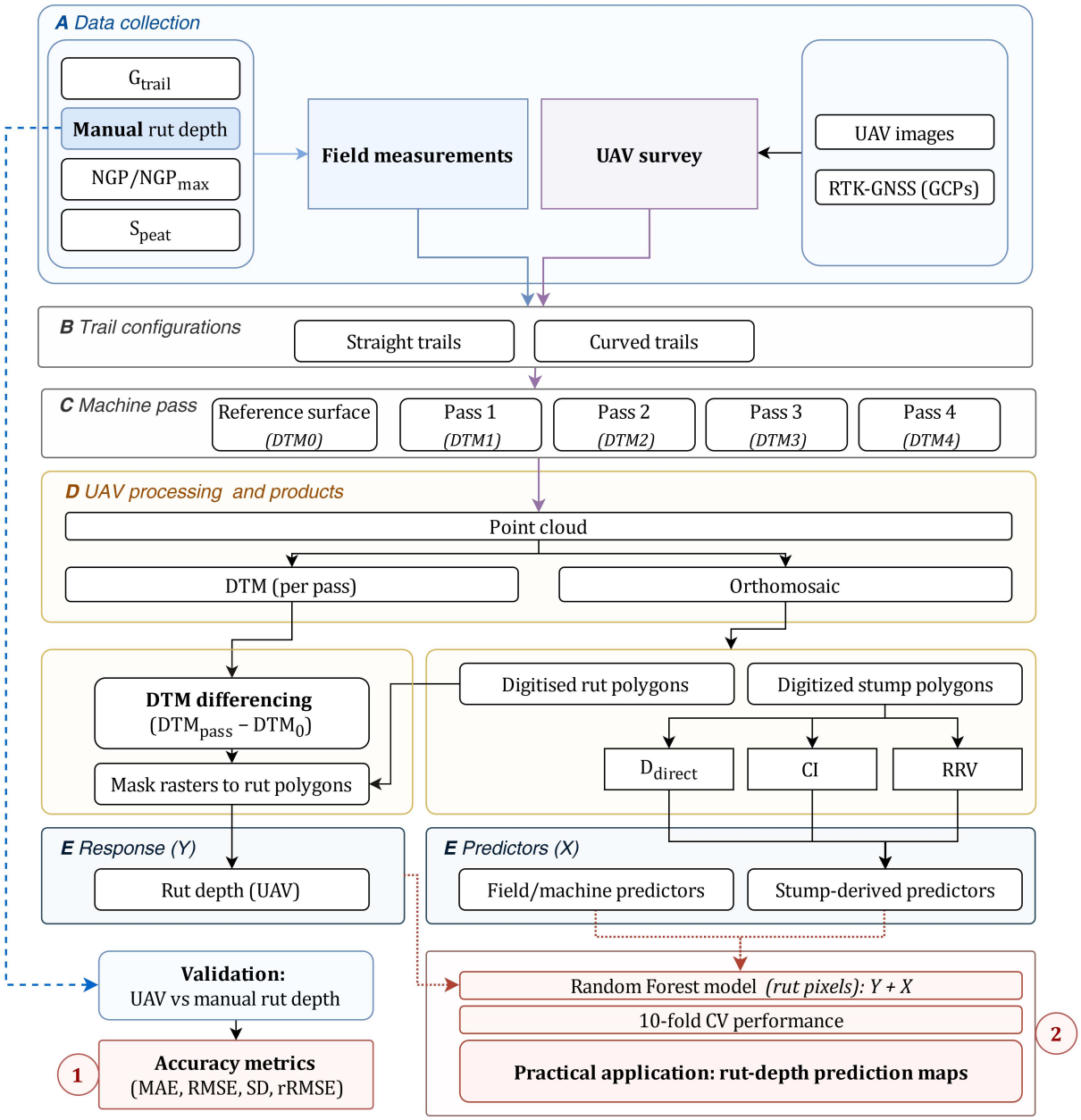


Figure 2.1: Workflow for rut-depth mapping and modelling. (A) Field and UAV data collection; (B) trail configurations; (C) machine passes; (D) UAV processing and terrain products (point clouds, DTMs, orthomosaics) and DTM differencing; (E) response (rut depth from DTM differencing) and predictor variables (field/machine and stump-derived). The workflow produces two main outcomes: (1) validation of UAV-derived rut depth against manual measurements using accuracy metrics and (2) Random Forest rut-depth prediction (10-fold cross-validation) and spatial prediction maps.

The region is characterised by a cold-temperate climate with pronounced seasonal variability. Winters are long, cold, and snowy, while summers are mild and short. The average annual temperature is around 2 to 4°C, with January the coldest (-10 to -15°C) and July the warmest (15 to 20°C). Rainfall is relatively evenly distributed throughout the year, averaging approximately 600-700 mm [121].

The study site is a mixed stand dominated by Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*). Peat depth in the area averages 149 cm, with a moisture content of

approximately 86 $\%_{wgt}$. The groundwater depth is relatively shallow, averaging approximately 29 cm.

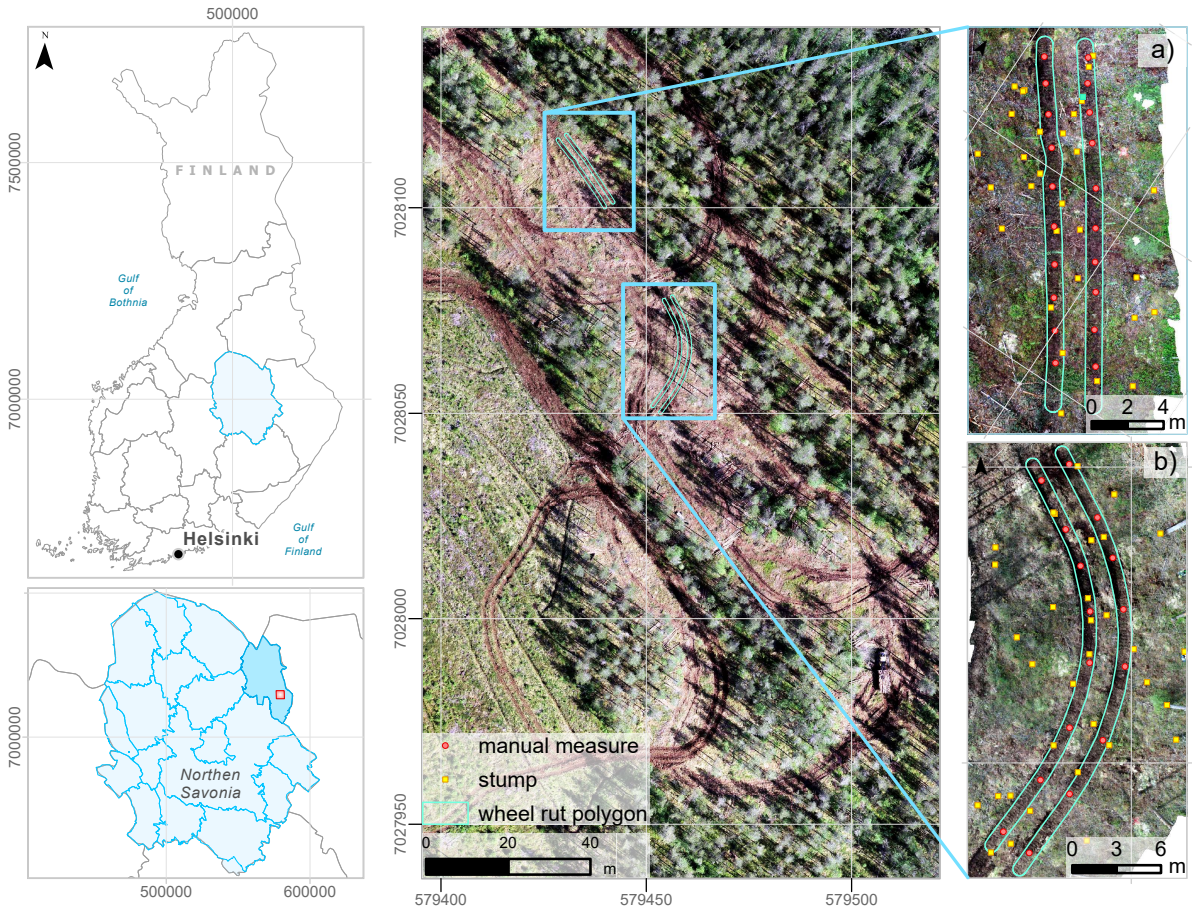


Figure 2.2: Study area in a peatland near Tiilikkala, Rautavaara, Northern Savonia, Finland. The figure displays straight and curved sections of Track 1 after the third machine pass. The left panel shows the site location, while the middle and right panels present orthomosaics of the track. Red dots indicate manual rut depth measurements along the wheel ruts (blue polygons), and green dots mark tree stumps. This setup was used consistently across all five tracks.

2.3.2 FIELD DATA COLLECTION

The study area comprises five test trails, each with two configurations - straight and curved. In this context, the term “trail” refers to the machine’s travel path, including both the left and right wheel ruts formed by the forwarder. The straight trails were 20 meters long and 4 meters wide, while the curved trails, also 4 meters wide, had a 20-meter radius and formed a 90° turn. The trails were carefully designed to enable forwarders to transition smoothly between straight and curved configurations without causing sharp turns or damaging the paths. To keep the soil and tree roots within the test trails undisturbed, the trails were prepared in advance using a harvester operating outside the test areas. Felled trees were processed away from the trails to avoid creating slash mats.

Five different forwarder models were tested, each equipped with a specific track configuration suited to the conditions. These forwarders were consistently used in both straight

and curved configurations. The forwarders included - the Ponsse Elk with a long wheelbase (LWB), the older generation Ponsse Elk with Fomatec tracks (EFwo), the older generation Ponsse Elk with Fomatec tracks and add-on track shoes (EFw), the Ponsse Buffalo with KOPA high flotation tracks (Kopa), and the Ponsse Elk 10W with Olofsfors mixed tracks (E10W). The data used in this study were initially collected as part of a previous investigation on rut formation in peatlands [105]. For detailed specifications of the machines used, please refer to the original study.

UAV-based data collection was conducted at the same experimental sites during the same 2019 field campaign, alongside manual rut depth measurements. UAV images were captured immediately before and after each machine pass, ensuring that all data sources reflect the same site conditions. As a result, no temporal changes needed to be considered between the different data collection methods.

The machine test mass includes the estimated load and the manufacturer-specified mass of the forwarder and tracks. The forwarder was loaded with approximately 5,100 kg of pulpwood, each piece averaging 4.56 meters in length. This load size, selected by the forwarder operator during the initial test, was chosen to enable three to four passes per trail with minimal risk of bogging down. Consequently, the forwarder completed the planned number of passes on each trail.

Rut depth was measured after each forwarder pass using a horizontal laser level and a surveyor's measuring rod. Measurements were taken at specific points along both the right and left wheel ruts, generally spaced every 2–3 meters across the trails. A total of 198 measurements were collected - 126 from straight sections and 72 from curved sections, with the difference reflecting variations in trail length. Since straight sections were longer, they naturally yielded more total measurements. These measurements were distributed across all test trails to ensure a representative assessment of rut depth variability in both configurations. To maintain consistency, measurement points were marked on the peat surface using spray paint. This ensured that we recorded new data at the exact locations throughout the passes.

Rut depth was measured at the deepest point, typically within a track-shoe depression in the peat. Additionally, peat moisture content (W_{peat}, %wgt) was recorded for each trail and configuration as a site-specific parameter. Shear modulus (G_{trail}, kPa), measured with a spiked shear vane [113], was also recorded at each measurement point. Nominal ground pressure (NGP) was determined based on the machine's technical specifications and depended on both the contact area and the machine's weight. The front and rear track sections were reported separately to account for variations in load distribution. For further analysis, the average (NGP_{mean}) and maximum (NGP_{max}) NGP values were used to characterise the machines' technical specifications. Three passes, each defined as a single traversal of the machine in one direction, were conducted on trails 1, 3, and 4, while trails 2 and 5 underwent four passes. A fourth pass was not performed at all sites because the peat had reached its strength limit.

2.3.3 UAV data collection and processing

Drone flights were conducted using the DJI Phantom 4 Pro (model: FC6310), flying at altitudes around 30-35 meters, capturing images over the same trails before the operations and immediately after each machine pass. A total of fifty ground control points (GCPs) were recorded using a Real-Time Kinematic Global Navigation Satellite System (RTK-GNSS) device. Each trail configuration had at least three GCPs, with an average of 56 images captured per area.

The RGB images and GCPs were processed in Agisoft Metashape Professional [122] to generate an initial point cloud dataset for each trail and its configurations. Agisoft Metashape combines Structure-from-Motion (SfM) and photogrammetric stereo-matching algorithms to reconstruct 3D models from unordered, overlapping imagery. The workflow included image import, alignment, georeferencing, and optimisation, followed by the generation of an initial point cloud. The GCPs were used to refine camera positions and improve georeferencing accuracy.

All point clouds were then exported in .LAS format for further processing. To ensure precise alignment across datasets, point clouds from each trail and pass were manually aligned in CloudCompare using the "*Align Two Clouds*" tool. At least seven point pairs were selected for point cloud registration, with the initial pre-operation dataset serving as the reference for aligning all subsequent point clouds. Fine registration using the iterative closest point (ICP) algorithm was subsequently applied to further improve dataset alignment [123].

To remove irrelevant vegetation, point clouds were classified using the cloth simulation filter (CSF) in CloudCompare [124]. The filtering parameters, such as cloth resolution (1.5), maximum iterations (1000), and classification threshold (0.5), were determined through trial and error to separate ground points from vegetation best while preserving terrain features. The aligned and classified point clouds were then reimported into Agisoft Metashape, where orthomosaics were generated for each trail and configuration. Only points classified as 'ground' were used to create a digital terrain model (DTM) at a resolution of 0.03 meters. Finally, all DTMs and orthomosaics were exported in GeoTIFF (.tif) format.

2.3.4 Rut depth analysis

To quantify elevation changes caused by each machine pass and assess the accuracy of UAV-based rut depth measurements, we used DTMs to compute elevation differences between passes. These changes were determined by subtracting the DTMs generated after each machine pass (DTM1, DTM2, DTM3, DTM4) from the initial DTM (DTM0), captured before operations for all trail configurations.

To evaluate the accuracy of the UAV depth measurements, we compared 198 manual reference points, as described in Section 2.3.2, with the mean DTM value within a 10 cm

buffer around each point. Accuracy was assessed separately for each pass and each trail to ensure consistency across different configurations. We then calculated error metrics, including the mean absolute error (MAE) (Eq. 2.1), root mean square error (RMSE) (Eq. 2.2), relative RMSE (RMSE%) (Eq. 2.3), and the standard deviation of differences (SD) (Eq. 2.4).

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2.1}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2.2}$$

$$\text{RMSE\%} = \left(\frac{\text{RMSE}}{\bar{y}}\right) \times 100 \tag{2.3}$$

$$\text{SD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left[(y_i - \hat{y}_i) - \overline{d}\right]^2} \tag{2.4}$$

In equations (2.1–2.4), y_i are the observed (manual) values, $\hat{y}_i$ are the predicted (UAV-derived) values, n is the number of observations, $\bar{y}$ is the mean of the observed values, and $\overline{d}$ is the mean difference $(y_i - \hat{y}_i)$. All statistical analyses and visualisations were performed using Python 3.12.

2.3.5 Tree stump influence and soil properties

2.3.5.1 Tree stump mapping

This study analysed tree stump positions and processed DTMs before and after each machine pass, along with manual measurements, to assess the influence of tree root systems on rutting. For each parameter described below, a corresponding raster map was created at a 0.03 m resolution, aligned with the DTM for each trail configuration. Wheel ruts in these areas were manually digitised, and all DTMs and other raster layers were clipped using these polygons. Only the pixels within these polygons were used in further analysis, ensuring that we focused specifically on machine-induced rutting effects.

To obtain precise locations, tree stumps within each trail and their configurations were manually identified using orthomosaics. Each orthomosaic was clipped using a 6-meter buffer around the trail centerline to define the mapping area. We assumed a 6-meter buffer was sufficient to capture all relevant stumps that could potentially influence rut formation within the affected soil zone. This area was then manually mapped to identify stump locations. A polygon was then created for each stump to define its boundaries, and the diameter was calculated by finding the largest distance between its vertices (Eq. 2.5).

$$d = \max\left(\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\right) \tag{2.5}$$

Equation (2.5) defines d as the largest distance between any two vertices, and (x_1, y_1) and (x_2, y_2) as the coordinates of any two vertices of the polygon. We also generated centroid points for further impact assessment. The summary of diameters across all sites and trails is presented in Fig. 2.6.

2.3.5.2 PEAT STRENGTH ESTIMATION

Peat soils exhibit highly variable bearing capacity, influenced by moisture content. In this study, peat strength refers to the unconfined compressive strength of peat and was estimated based on existing guidelines from plate loading tests, which reflect the unconfined compressive strength of peat under deformation conditions [125].

In dry conditions, peat strength typically ranges from 40 to 70 kPa, while in moist conditions, it decreases to between 10 and 40 kPa [125]. The peat's moisture content on the site averaged 86% (%wgt), ranging from 84% to 89% based on field measurements. To align the peat strength with the specific moisture levels of each trail, we estimated the peat strength (S_{peat}) for each pixel using the following formula:

$$S_{\text{peat}} = 40 - 3 \times (W_{\text{peat}} - 85) \tag{2.6}$$

where S_{peat} is the estimated peat strength (kPa), and W_{peat} is the peat moisture content ($\%_{\text{wgt}}$). In this formula, the base strength is 40 kPa at a moisture level of 85%. The strength increases to approximately 42 kPa at the lowest recorded moisture level of 84%, while it decreases to approximately 27 kPa at the highest recorded moisture level of 89%.

2.3.5.3 DISTANCE TO THE STUMP

The influence of stumps on soil stability diminishes with distance [114]. To quantify this effect, we used an exponential decay model to represent the cumulative influence of multiple stumps (Fig. 2.3). While exponential decay models are widely used in environmental and geomechanical studies to describe processes such as root reinforcement decay and soil production [126], to our knowledge, they have not yet been explicitly applied to understanding the role of stumps in rut formation.

The modelled influence follows an exponential decline, with the effect strongest near the stump and gradually decreasing with distance (Eq. 2.7).

$$I(d) = e^{-k \cdot d} \tag{2.7}$$

$I(d)$ is the influence at a distance d, d is the distance from the centre of the stump to the centre of the pixel being evaluated in m, and k is the decay constant that controls how quickly the influence diminishes. A higher k value indicates a rapid decline in influence, while a lower k value means the influence extends over a broader area.

We determined the decay constant ($k = 0.3$) by evaluating a range of candidate values ($k = 0.1, 0.2, 0.3, 0.4, 0.5, 1.0$) in an iterative manner. The selected value was chosen to achieve a realistic balance between the influence range and decay rate, ensuring that the influence extends spatially in a reasonable manner without exaggerating its effect.

To determine the distance from the stump to the centre of each pixel d, we used the Euclidean distance equation (Eq. 2.8).

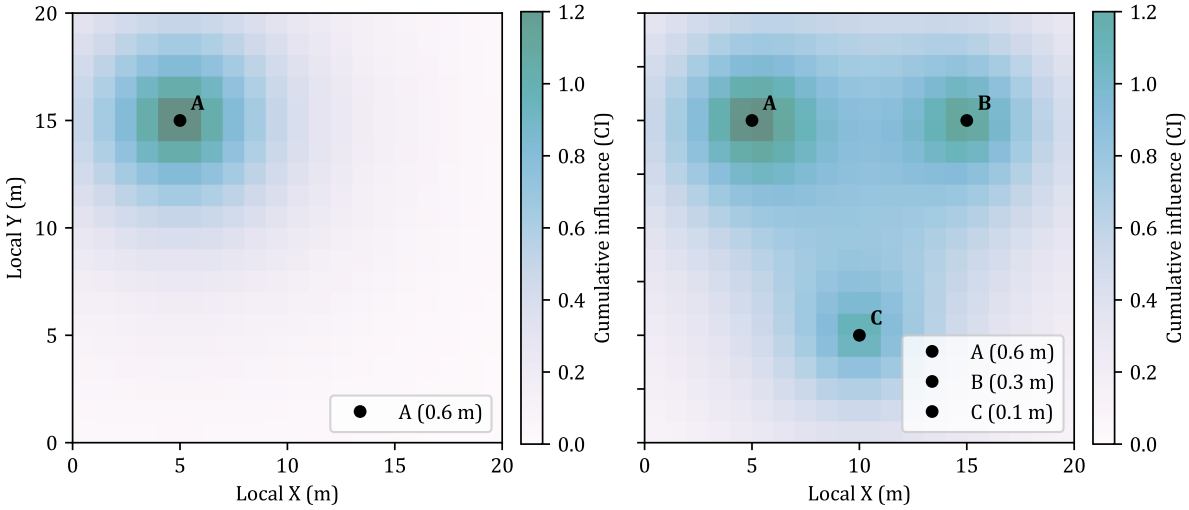
$$\text{distance} = \sqrt{(x_{\text{pixel}} - x_{\text{stump}})^2 + (y_{\text{pixel}} - y_{\text{stump}})^2} \tag{2.8}$$

In this case, x_{stump} and y_{stump} are the coordinates of the stump's centroid, $x_{\text{pixel}}, y_{\text{pixel}}$ are the coordinates of the centre of each pixel in the raster.

Given that larger stumps have a more substantial influence due to more extensive root systems, we adjusted the function to include stump diameter as a scaling factor (Eq. 2.9).

$$I(d, D) = \left(1 + \alpha \cdot \frac{D}{D_{\text{ave}}}\right) \cdot e^{-k \cdot d} \tag{2.9}$$

D is the stump diameter, D_{ave} is the average diameter of all stumps, and α is a scaling constant controlling the effect of stump diameter on influence.



(a) Influence of stump A (b) Cumulative influence

Figure 2.3: Visualisation of distance-based stump influence. (a) Influence of a single stump (A) modelled using an exponential decay function. (b) Cumulative influence using the same exponential decay function from multiple nearby stumps (A, B, C).

To account for overlapping influences from multiple stumps, the cumulative influence (CI) at each pixel (Fig. 2.3) was calculated as the sum of the adjusted influences from all nearby stumps (Eq. 2.10). This cumulative effect aims to replicate actual environmental conditions, where multiple tree root systems often interact, creating an interlinked network that improves soil stability [127].

$$CI = \sum_{\text{stumps}} I(d, D) \tag{2.10}$$

In Eq. (2.10), $I(d, D)$ is the influence of an individual stump at a distance d from the pixel, scaled by the stump's diameter D. CI represents the cumulative influence at each pixel, calculated by summing contributions from all stumps. Stumps with $D > D_{\text{ave}}$ have a scaling factor greater than 1, increasing their influence; those with $D < D_{\text{ave}}$ have a scaling factor less than 1, reducing their impact. This summation accounts for the overlap of the root systems and their combined effects.

2.3.5.4 Root reinforcement modeling

To estimate root reinforcement in our study, we used a simplified interpolation and curve-fitting method based on the findings of [114]. That study reported maximum reinforcement values for Norway spruce (*Picea abies*) at specific distances from the tree trunk, grouped by stump diameters of 20, 40, and 60 cm. We expanded on this data by interpolating reinforcement values at 50 cm intervals. Using the interpolated values, we applied a reverse sigmoidal function (Eq. 2.11) to create smooth curves for each stump diameter group. The reverse sigmoidal function, in this case, captures the natural decline of root support, being most robust near the stump and gradually declining outward. This resulted in specific equations for each group (Eq. 2.11, Table 2.1). The general function is defined as Eq. 2.11.

$$RRV(x) = \frac{L}{1 + e^{-k(x - x_0)}} \tag{2.11}$$

RRV is the root reinforcement value (kPa) at a given distance x from the stump, where x is measured in m. L is the maximum reinforcement level derived from the data (kPa), k controls the steepness of the curve, x_0 is the distance at which the curve has an inflection point, x is the distance from the stump to the pixel being evaluated.

These coefficients were then applied iteratively to each digitised stump across the study area, considering the stump diameter, the distance from the stump centre to the centre of each pixel (Eq. 2.8), and the coordinates of each pixel.

Table 2.1: Logistic curve coefficients for different stump diameters.

Diameter class (cm)	L (kPa)	k	x_0
20	21.09	-1.006	2.61
40	45.16	-1.046	3.74
60	79.55	-1.016	4.01

Specific parameters (Table 2.1) were used to compute the reinforcement value for each pixel based on its distance from the stump, following a sigmoidal decay curve. The influ-

ence radius was guided by the approach described in [114], which relates stump diameter to the extent of root reinforcement, with distances indicating where reinforcement approaches zero (RRV = 0). For areas where the reinforcement effect of multiple stumps overlapped, we summed the reinforcement values to account for the combined effect. This was done by iterating over each stump, applying the logistic function within its influence radius, and accumulating the values in an RRV raster. The resulting RRV raster values were normalised to the 0–100 scale using Eq. (2.12), yielding a root reinforcement map that reflects the combined spatial influence of all stumps across the study area.

$$R_{\text{norm}} = 100 \cdot \frac{R_{\text{final}} - R_{\text{min}}}{R_{\text{max}} - R_{\text{min}}} \tag{2.12}$$

R_{final} represents the summed reinforcement values, and R_{min} and R_{max} are the minimum and maximum reinforcement values across the study area.

2.3.6 Statistical analysis

To evaluate how tree stumps influence rut depth, we categorised CI values within each trail and configuration into five groups: ”very high”, ”high”, ”moderate”, ”moderately low”, and ”low”. Areas classified as ”very high” contained dense, large stumps, exerting a strong influence, whereas ”low” areas had fewer stumps and less impact. The categorisation was based on the distribution of CI values, which were grouped using natural breaks identified in the histogram and kernel density estimation (KDE) plots. These breakpoints were chosen to represent shifts in the data, ensuring each group captured a distinct range of influence.

The data were first tested for normality using the Shapiro-Wilk test across straight and curved configurations. Since the results indicated a non-normal distribution, we used the Kruskal-Wallis test to assess whether rut depth significantly differed among distance groups within each pass and trail configuration. If $p < 0.05$, then at least one distance group had a significantly different rut depth. If $p > 0.05$, we concluded that rut depth did not vary substantially across distance groups within that configuration. Due to the large and uneven sample sizes across distance groups, we used balanced random subsampling to ensure fair statistical comparisons and mitigate potential large-sample biases. To additionally assess the stability of our samples, we repeated the Kruskal-Wallis test across five independent iterations, each using a new random subsample from the dataset.

To determine the effect size of rut depth variations, we computed epsilon-squared (ε^2) (Eq. 2.13) as a measure of effect size for the Kruskal-Wallis test. This metric helps determine the proportion of variance in rut depth explained by distance groups, with values above 0.06 indicating a moderate effect and values above 0.14 (Table 2.2) suggesting a strong influence of stump presence on rut depth formation [128, 129].

$$\varepsilon^2 = \frac{H}{\frac{N^2-1}{N+1}} \tag{2.13}$$

where, H is the Kruskal-Wallis test statistic, and N is the total number of observations.

A small effect size might suggest a significant but weak influence of stump proximity on rut depth. In contrast, a large effect size would indicate that stump proximity has a substantially greater impact on rut formation.

Table 2.2: Effect size (ε^2) interpretation thresholds for the Kruskal-Wallis test.

Effect size	**Threshold (ε^2)**
Small	$\varepsilon^2 \geq 0.01$ and < 0.06
Medium	$\varepsilon^2 \geq 0.06$ and < 0.14
Large	$\varepsilon^2 \geq 0.14$

2.3.7 MACHINE-LEARNING FOR DEPTH PREDICTION

To evaluate how well rut depth can be predicted using soil properties, stump influence, and environmental factors, we implemented a Random Forest (RF) model. Machine learning algorithms identify complex spatial and non-linear relationships between inputs and targets that traditional statistical methods often fail to capture. For instance, RF constructs multiple decision trees and combines their outputs for more reliable predictions [130]. We employed an RF model to predict rut depth based on soil properties, stump presence, and other factors. The model used only pixels with DTM-derived depth values below zero as the dependent variable (Y), while soil, environmental, and machine parameters served as independent variables (X) (Table 2.3).

The selection of RF was based on the data distribution and the complexity of the predictor variables. Kolmogorov-Smirnov tests and skewness analyses indicated that the key variables (CI, RRV, and SM) were highly skewed, thereby rejecting the normality assumption required for linear regression. Additionally, RRV and CI follow logistic and exponential decay, both of which introduce nonlinearity and interaction effects.

The model was trained separately for each pass over various datasets, including (a) the full or combined dataset as well as separate datasets for (b) straight and (c) curved trails. We tuned the RF hyperparameters to improve model accuracy using RandomizedSearchCV [131]. The final configuration used was $n_{\text{estimators}} = 200$, $max_{\text{depth}} = \text{None}$, $min_{\text{samples\_split}} = 5$, and $min_{\text{samples\_leaf}} = 1$.

Feature importance scores were initially determined using the RF model, after which recursive feature elimination (RFE) was utilised to refine the model for both straight and curved trail configurations, as well as for the combined dataset. This process iteratively removes the least essential features while maintaining predictive performance. The final model retained the top three features ranked by importance. A separate RF model was trained for each configuration and pass.

The data were split into 70% training and 30% testing sets for each pass, and the model's performance was evaluated using R-squared (Eq. 2.14) scores on both sets.

Table 2.3: List of dependent (Y) and independent (X) variables

Variable	Description	Units
Y	Rut depth (RD)	m
X_1	Root deinforcement value (RRV)	arbitrary units
X_2	Cumulative distance from the stump (CI)	arbitrary units
X_3	Moisture-dependent peat strength (S_{peat})	kPa
X_4	Shear modulus (G_{trail})	kPa
X_5	Maximum NGP (NGP_{max})	kPa
X_6	Mean NGP (NGP_{mean})	kPa

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{2.14}$$

where, n is the number of data points, y_i is the actual value of the i-th data point, $\hat{y}_i$ is the predicted value of the i-th data point; $\bar{y}$ is the mean of the actual values ($\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$), $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ is the residual sum of squares (SSR), $\sum_{i=1}^{n} (y_i - \bar{y})^2$ is the total sum of squares (SST).

Additionally, ten-fold cross-validation was performed on the complete dataset for each pass to assess model stability. While the same data points may appear in both training and validation sets across different folds, each fold was strictly separated, ensuring that no data point was used for both training and testing within the same fold. The best-performing model was selected based on the highest average R^2 (Eq. 2.14).

To demonstrate the model's performance, we applied the best-performing Random Forest configuration to a synthetic spatial dataset comprising simulated shear modulus values and randomised tree-stump distributions. Shear modulus values were generated by randomly placing control points across the area and interpolating them over a grid using cubic interpolation to create a continuous surface. Tree stumps were also randomly distributed, with diameters ranging from 10 to 40 cm and a minimum spacing of 2 m between them (60 stumps in total). To simulate more natural clustering, each stump's diameter was averaged with those of nearby stumps. The key predictor variables identified in the model analysis were used to generate corresponding pixel-wise input datasets. The trained model was then applied to this synthetic area, and the predictions were visualised as a contour map. The input raster grid resolution was initially set to 0.03 m and subsequently resampled to 0.5 × 0.5 m for visualisation.

2.4 Results

2.4.1 Rut depth analysis

The average rut depth across all trails and configurations was 15.5 cm for manual measurements and 12.5 cm for UAV measurements, with standard deviations of 7.6 cm and 8.3 cm, respectively. Both datasets showed variation in rut depth across trails and configurations, with differences influenced by the number of passes (Table 2.4).

Table 2.4: Summary statistics of manual and UAV measurements across passes. Each value represents the overall statistic, with values in parentheses indicating the straight (first) and curved (second) configurations.

	Mean (manual)	**Mean** (UAV)	**SD** (manual)	**SD** (UAV)	**MAE** (cm)	**RMSE** (cm)	**RMSE** (%)
Pass 1	11.1 (11.3, 10.7)	10.6 (10.9, 9.9)	5.9 (6.4, 4.8)	5.3 (5.9, 4.0)	2.2 (1.8, 2.8)	3.0 (2.4, 3.6)	26.5 (21.4, 34.1)
Pass 2	15.0 (14.9, 15.1)	14.1 (13.5, 15.0)	6.6 (7.2, 5.5)	6.0 (6.4, 5.0)	2.3 (2.0, 2.7)	3.1 (2.6, 3.8)	20.8 (17.5, 25.3)
Pass 3	19.3 (19.2, 19.5)	17.6 (17.6, 17.7)	7.3 (7.8, 6.3)	7.0 (7.7, 5.7)	2.7 (2.4, 3.1)	3.5 (3.0, 4.2)	18.3 (15.9, 21.8)
Pass 4	20.5 (20.6, 20.2)	19.8 (19.8, 19.9)	7.6 (8.1, 6.6)	6.8 (6.9, 6.7)	3.3 (2.7, 4.2)	4.2 (3.4, 5.2)	20.3 (16.5, 25.7)
Total	15.5	12.5	7.6	8.3	2.5	3.3	24.9

For manual measurements, rut depth ranged from –8 cm (N = 4) to 41 cm, with negative values due to upwelling or measurement difficulties near stumps. UAV measurements followed a similar trend but ranged from 0 cm to 40 cm, showing slightly lower values for deeper ruts, particularly on curved trails (Fig. 2.4). Across passes, UAV consistently measured lower rut depths than manual measurements. In Pass 1, the UAV underestimated by 0.5 cm (10.6 cm vs. 11.1 cm), with the gap increasing slightly in later passes. In Pass 4, the UAV measured 19.8 cm compared to 20.5 cm manually, maintaining a consistent underestimation trend. Measurement variability also increased as ruts deepened. SD remained similar in early passes (5.9 cm manual and 5.3 cm UAV in Pass 1) but showed a slight decrease in UAV measurements in later passes (7.6 cm manual and 6.8 cm UAV in Pass 4). The UAV underestimation was more consistent along straight trails, but greater variability was observed in curved trails, where the SD was 8.1 cm for manual measurements in Pass 4, compared to 6.7 cm for UAV.

Rut depth increased from Pass 1 to Pass 4, as expected, with each pass further compacting the soil and increasing rut formation. This effect was more pronounced on curved trails, where lateral forces exerted by the machine contributed to deeper ruts by both compacting and displacing the soil. These trends were consistent with findings in the same study area [105]. The boxplots in Fig. 2.4 illustrate this pattern, showing that straight trails exhibited

lower rut depths than curved trails in both manual (Fig. 2.4a) and UAV-derived (Fig. 2.4b) measurements.

While both methods captured the overall trend of rut formation, differences in accuracy were observed. Manual measurements (Fig. 2.4a) showed a broader distribution of rut depths, particularly in the earlier passes, while UAV measurements (Fig. 2.4b) followed a similar pattern but with consistently lower values.

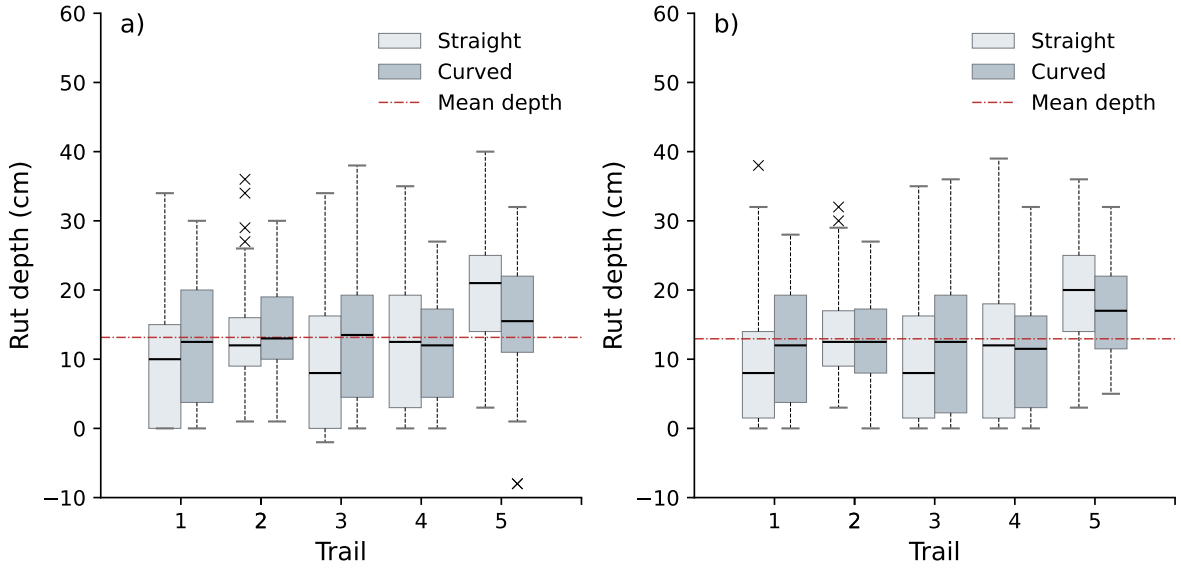


Figure 2.4: Descriptive statistics of manual and UAV rut depth measurements. (a) Boxplot of manual rut depth measurements by trail type (straight and curved). The box shows the 1st (25%) and 3rd (75%) quartiles, while the whiskers represent data within 1.5 times the interquartile range (IQR). The red dashed line represents the mean rut depth of 13.2 cm across all manual measurements. (b) Boxplot of UAV rut depth measurements by trail type (straight and curved). The same style as (a), with the red dashed line representing the mean rut depth for UAV measurements of 12.9 cm across all UAV measurements.

Measurement errors increased as passes progressed. The MAE increased from 2.2 cm in Pass 1 to 3.3 cm in Pass 4, while RMSE increased from 3.0 cm to 4.2 cm (Fig. 2.5). RMSE was highest in curved trails, reaching 5.2 cm in Pass 4 compared to 3.4 cm in straight trails, highlighting the UAV's difficulty in capturing the full depth of irregular ruts. SD for both manual and UAV measurements increased with each pass, showing greater variability in rut formation as deeper ruts developed due to soil displacement and compaction. Several outliers (>30 cm) were observed, primarily on curved trails, where machine-induced lateral displacement resulted in more uneven rut profiles.

2.4.2 Tree stump influence and soil properties

Tree stump diameters ranged from 5.4 cm to 30.0 cm, with fewer stumps at the extreme ends of this range. The average diameter was 22.5 cm, with a standard deviation of 8.1 cm, indicating moderate variation. In total, 269 stumps were analysed. Fig. 2.6 shows the

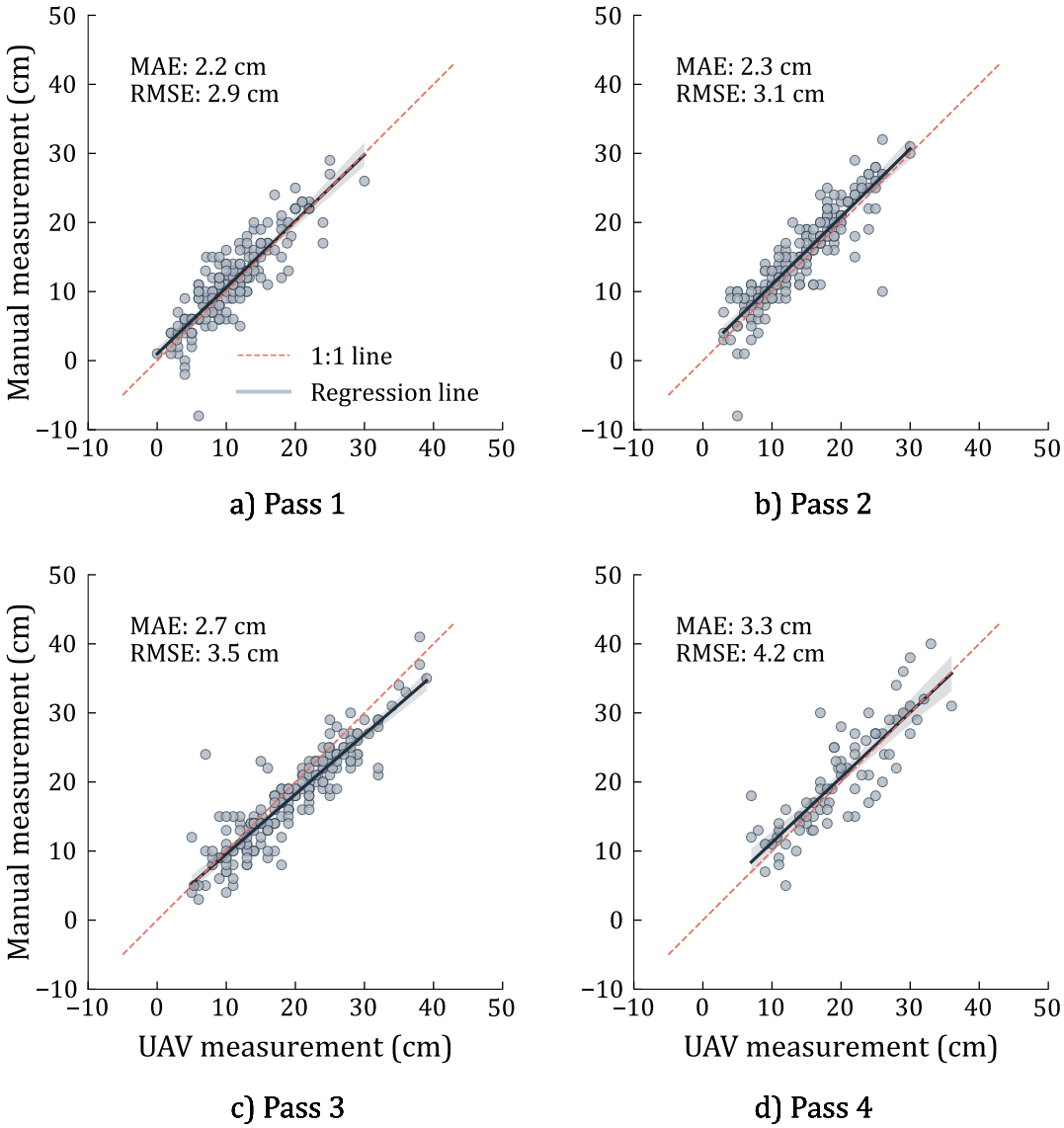


Figure 2.5: Scatterplot of manual and UAV measurements for Pass 1 to 4 (a–d). The blue line represents the regression line, while the red dashed line indicates the 1:1 reference line. The mean absolute error (MAE) and root mean square error (RMSE) are provided in cm for each pass.

distribution of stump diameters across different trails and configurations with consistent median values. No significant differences were observed between straight and curved trail configurations, and the overall average diameter remained stable across all trails.

Peat moisture content averaged 86% across all trails, ranging from 84% to 89%, with higher moisture levels observed on curved trails. The estimated peat strength (Table 2.5) varied accordingly, ranging from 35.0 to 42.6 kPa on straight trails and 27.0 to 37.7 kPa on curves, indicating weaker soil conditions and higher susceptibility to rutting in curved sections. The shear modulus of the peatland top layer exhibited significant variability, with values ranging from 30.8 to 383.4 kPa on straight trails, while curved trails showed considerably lower stiffness, ranging from 23.9 to 138.3 kPa (Table 2.5).

To illustrate how stump presence and terrain changes interact spatially, Fig. 2.7 presents UAV-derived elevation differences (a–c), calculated as the difference between the

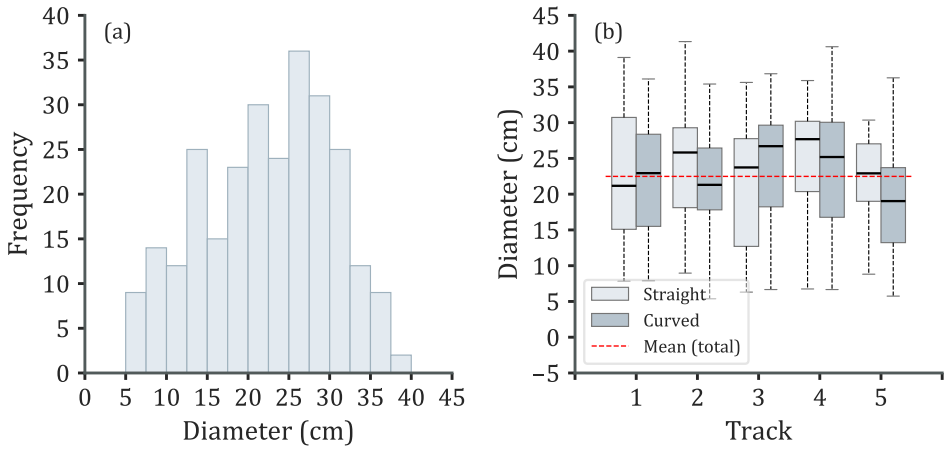


Figure 2.6: Descriptive statistics of tree stump diameters (cm) from digitised polygons. The histogram (a) shows the diameter distribution in cm. The boxplot (b) displays average diameters by trail and configuration (straight or curved), with boxes representing the interquartile range (IQR) and whiskers extending to the 95% quantiles. The line inside each box is the median, and the red dashed line indicates the overall mean diameter.

Table 2.5: Summary of peat moisture, peat strength, and shear modulus by configuration.

Configuration	Peat moisture (%)			Peat strength (kPa)			Shear modulus (kPa)		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Straight	84.1	86.7	85.2	35.0	42.6	39.3	35.2	383.4	86.3
Curved	85.7	89.3	87.0	27.0	37.8	34.0	23.9	138.3	51.8

post-pass DTM and the pre-operation DTM, alongside the cumulative stump influence map (d) for one representative straight trail. Areas with stronger cumulative influence, typically near larger or clustered stumps, tend to align with zones of reduced rutting, suggesting a local stabilising effect.

2.4.3 Statistical analysis

To assess how stump presence influenced rut formation, rut depth values were analysed across the five CI groups (very high, high, moderate, mod-low, and low) (Table 2.6). The Shapiro-Wilk test indicated a non-normal distribution for both straight ($W = 0.97$, $p < .05$) and curved ($W = 0.93$, $p < .05$) configurations. Therefore, the Kruskal-Wallis test was performed to compare the rut depth across influence groups. Results showed significant differences in rut depth among influence groups across all passes and configurations ($p < 0.05$) (Table 2.7). To ensure fair statistical comparisons and mitigate bias arising from unequal sample sizes, balanced random subsampling ($n = 5000$ per group) was used. The Kruskal-Wallis test was repeated across five independent iterations using different random subsamples, confirming that results remained stable. These results are presented in Table 2.7.

The H-statistic from the Kruskal-Wallis test measures rank-based differences between

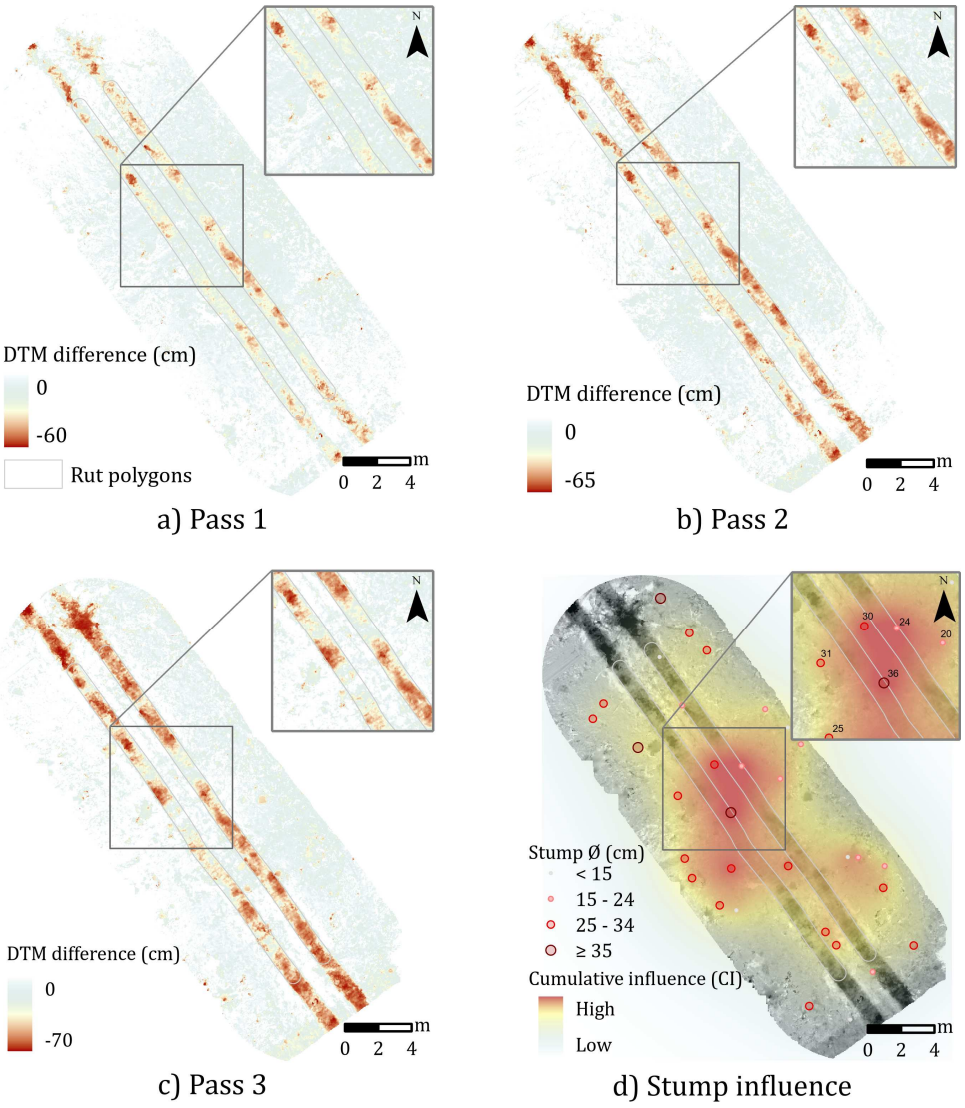


Figure 2.7: Elevation change maps derived from UAV-based digital terrain models (DTMs) for one experimental trail (straight configuration). Panels (a–c) show elevation differences after Passes 1, 2, and 3, respectively, calculated as DTM after each pass minus the initial pre-operation DTM. Panel (d) shows the cumulative stump influence (CI), computed from stump diameters and proximity using the methodology described above.

groups, with higher H-values indicating bigger differences among the CI groups. For straight trails, the H-statistics ranged from 94.61 (Pass 1) to 280.79 (Pass 4), with Pass 4 showing the highest value, indicating bigger differences between the groups. This suggests that, during Pass 4, rut depth differed significantly among CI groups, indicating that stump presence was associated with spatial variation in rutting intensity. For curved trails, the H-statistics ranged from 74.71 (Pass 1) to 524.38 (Pass 4), with similarly significant differences in rut depth across influence groups, especially in Pass 4. This indicates that stump presence affects rut depth across both trail types, but the effect is stronger in curved sections, possibly due to terrain differences.

The epsilon-squared values indicate that stump presence has a moderate to large effect on rut depth in curved trails, with medium effects observed in straight trails, especially in Pass 4. For straight trails, the ε^2 values ranged from 0.04 to 0.20, indicating a small to

Table 2.6: Summary statistics of rut depth by cumulative influence (CI) groups and configurations. Values represent median, mean, and standard deviation (SD) in centimetres.

Influence group	Configuration	Median	Mean	SD
Low	Straight	14	15	9
	Curved	18	19	11
Mod-Low	Straight	14	15	9
	Curved	16	18	11
Moderate	Straight	14	15	8
	Curved	14	15	10
High	Straight	13	14	8
	Curved	12	13	9
Very High	Straight	9	12	9
	Curved	9	10	8

Table 2.7: Kruskal-Wallis test for depth variations among cumulative influence (CI) groups by trail configuration and pass. The Epsilon-squared (ε^2) column is colour-coded to reflect effect strength: light grey for small, light blue for medium, and deeper blue for large effects, as defined in Table 2.2.

Configuration	Pass	H-statistic	p-value	ε^2
Straight	1	94.6	<0.05	0.06
	2	219.0	<0.05	0.07
	3	119.2	<0.05	0.04
	4	280.8	<0.05	0.20
Curved	1	74.7	<0.05	0.02
	2	397.2	<0.05	0.10
	3	334.9	<0.05	0.09
	4	524.4	<0.05	0.32
Overall		**1701.27**	**<0.05**	**0.07**

medium effect size. Pass 4 had the highest effect size ($\varepsilon^2 = 0.20$), suggesting that the influence of stumps on rut depth was most pronounced in this pass. For curved trails, the ε^2 values ranged from 0.02 to 0.32, showing a similar pattern. Pass 4 again had the highest effect size ($\varepsilon^2 = 0.32$), which is considered a medium to significant effect, indicating a substantial impact of stump presence on rut formation in curved trails.

The results are further visualised in Fig. 2.8, which presents boxplots of rut depth across influence groups for each pass and configuration. The plots show a consistent trend: rut depth is lowest in areas with high stump presence (very high CI) and increases with distance from stumps (low CI).

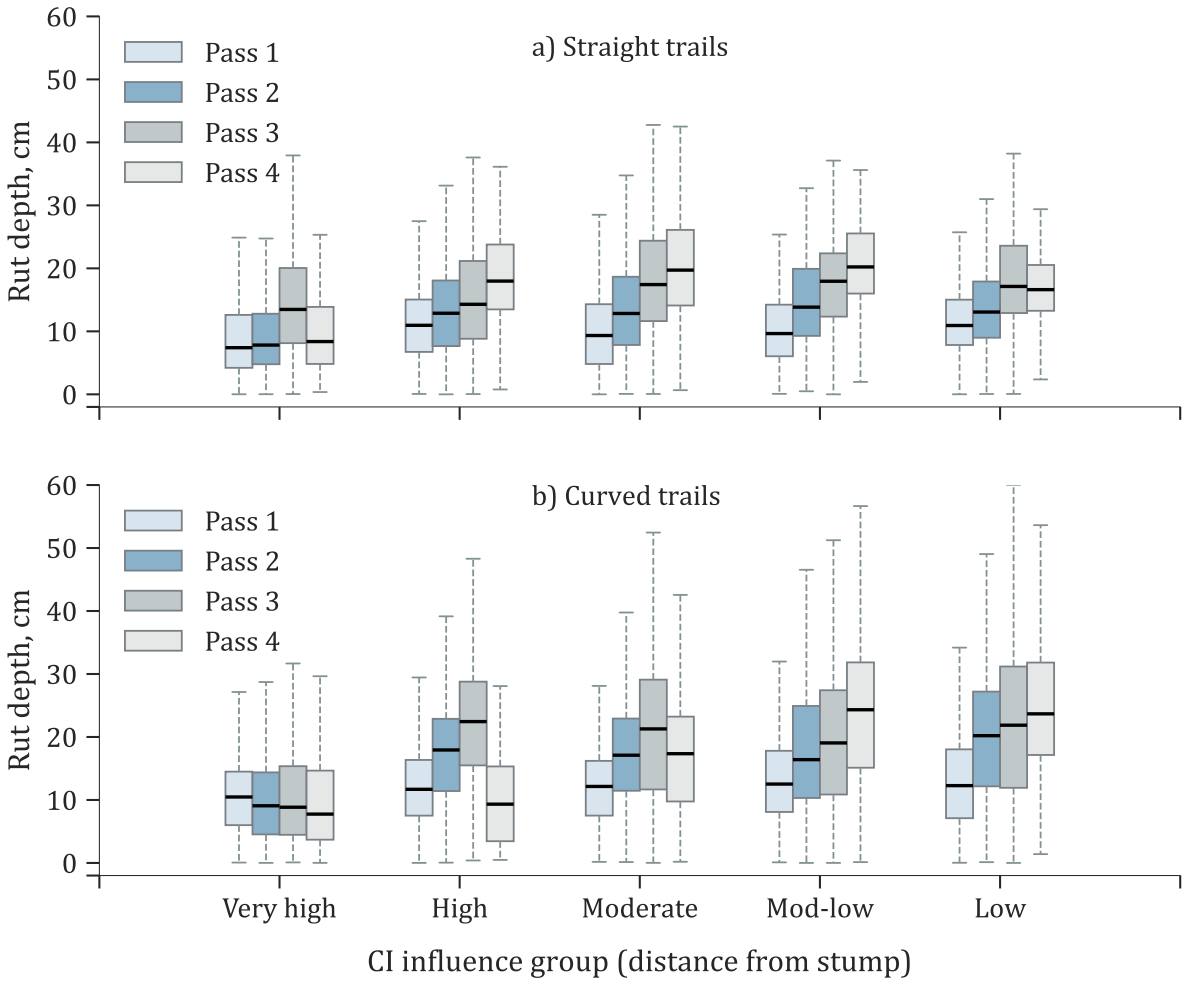


Figure 2.8: Boxplots showing rut depth (in cm) distribution across CI influence groups for each pass and trail configuration. The upper plot represents the straight trail.

2.4.4 ML for depth prediction

After training the model to predict rut depth, the RF models identified and ranked the most influential predictor variables. The results show that cumulative stump influence, root reinforcement value, and shear modulus were the top three predictors across all trail configurations (combined, straight, and curved datasets). These features were consistently ranked highest by the RF model and selected by RFE. In contrast, moisture-dependent peat strength, along with mean and maximum NGP, had minimal contribution across all configurations (Fig. 2.9).

A detailed analysis of feature significance across various trail setups shows distinct patterns. In the combined dataset (Fig. 2.9a), RRV (0.394) and CI (0.384) contributed almost equally to rut depth prediction, with RRV slightly outperforming CI. G_{trail} was the third most important feature but had a much lower score (0.109), indicating a weaker influence. For straight trails (Fig. 2.9b), RRV (0.429) and CI (0.419) recorded their highest importance scores, reinforcing their dominant role in rut formation. In contrast, G_{trail} had minimal influence (0.094), while the remaining features had negligible impact. In curved trails (Fig. 2.9c), RRV remained the most influential predictor (0.421), but CI's importance

slightly declined (0.358). Interestingly, G_{trail} played a greater role (0.146), suggesting that variations in shear modulus may be more influential in curved terrain, where soil resistance dynamics are likely to differ from those in straight trails.

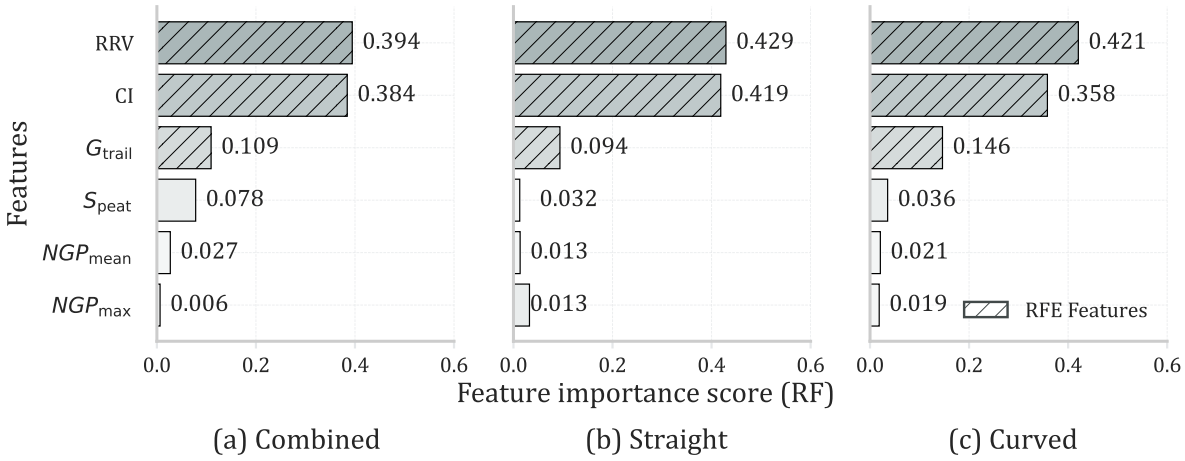


Figure 2.9: Feature importance results using the Random Forest (RF) model, where the analysed features are CI (cumulative influence of the stump based on distance from the stump), RRV (root reinforcement value), G_{trail} (shear modulus in kPa), S_{peat} (moisture-dependent peat strength), NGP (mean nominal ground pressure from front and rear in kPa), and Max NGP (maximum nominal ground pressure in kPa). The graph is divided by trail configuration into three categories: combined, straight, and curved, with the line pattern indicating the three most important features identified through Recursive Feature Elimination (RFE).

It is important to note that both RRV and CI were derived using similar approaches, as they both incorporate stump diameter and distance to the stump as key inputs. Due to their conceptual similarities, they exhibited a strong correlation, with R^2 = 0.76. However, they capture different aspects of the influence of stump roots on soil stability. Therefore, both were retained for further model training, as feature importance analysis consistently ranked them among the top predictors across all setups. Model performance tests showed that removing CI led to a 78.6% reduction in predictive accuracy, whereas removing RRV led to an 80.8% reduction, highlighting their contributions to model performance. RF performs best when it can leverage non-linear relationships and feature interactions. Removing one feature limits the model's capacity to detect threshold effects in RRV, which indicate root strength, and to account for distance-decay effects in CI, where influence decreases with distance. This means that features provide important information to the model. Therefore, RRV, CI, and G_{trail} were used for further model training (Fig. 2.9).

The dominance of RRV and CI across all configurations suggests that stump presence plays an essential role in soil stability and rut formation. The more substantial influence of G_{trail} in curved trails implies that shear modulus variations have a greater effect on rut formation when the machine path is nonlinear. The negligible impact of peat strength, NGP_{mean}, and NGP_{max} suggests that nominal ground pressure alone may not sufficiently

explain rut formation, reinforcing the need to include root-related features for more accurate predictions.

Table 2.8: RF results presented as average performance metrics from 10-fold cross-validation by configuration and pass, with training and testing dataset sizes indicated in thousands of pixels (Kpx).

Configuration	Pass	Train size (Kpx)	Test size (Kpx)	R^2 (train)	R^2 (test)
Straight	1	160.5	40.1	0.96	0.69
	2	160.6	40.1	0.96	0.74
	3	172.7	43.1	0.97	0.77
	4	74.2	18.5	0.96	0.74
Curved	1	238.6	59.6	0.96	0.75
	2	226.9	56.7	0.97	0.81
	3	200.6	50.1	0.98	0.85
	4	84.3	21.1	0.97	0.82
Combined	1	399.1	99.7	0.96	0.73
	2	387.5	96.8	0.97	0.80
	3	373.3	93.3	0.98	0.83
	4	158.6	39.6	0.97	0.79

To further validate the RF model, 10-fold cross-validation was performed across different configurations (Fig. 2.10), and the results are summarised in Table 2.8. The model demonstrated high predictive capability, with training R^2 values consistently above 0.96 across all configurations, and test R^2 values varying across different trail types. The model achieved a test R^2 ranging from 0.73 to 0.85 in the combined dataset, indicating good generalisation across all trail types. On straight trails, lower test R^2 values (0.69-0.77) indicate that the model performed worse at predicting rut depth than on curved trails. The highest test R^2 values (0.75 to 0.85) were observed in curved trails, indicating better predictive accuracy in these conditions (Fig. 2.10).

The lower R^2 values in straight trails may be due to repeated passes over the same trail, leading to progressive displacement and surface-layer compaction, making rut formation less sensitive to individual predictor variables. Repeated traffic in straight paths compresses the same area, leading to increasingly uniform soil resistance and reduced variability in rut depth. Conversely, curved paths exhibit greater variability due to uneven soil compaction and lateral displacement pressures. On curved trails, the turning motion of the steel trails can induce lateral displacement and localised shear stresses in the soil, particularly when differential lock is engaged. This creates uneven loading and increased soil deformation, contributing to greater variability in rut development, which the model captures more effectively. This can result in greater variability in rut development,

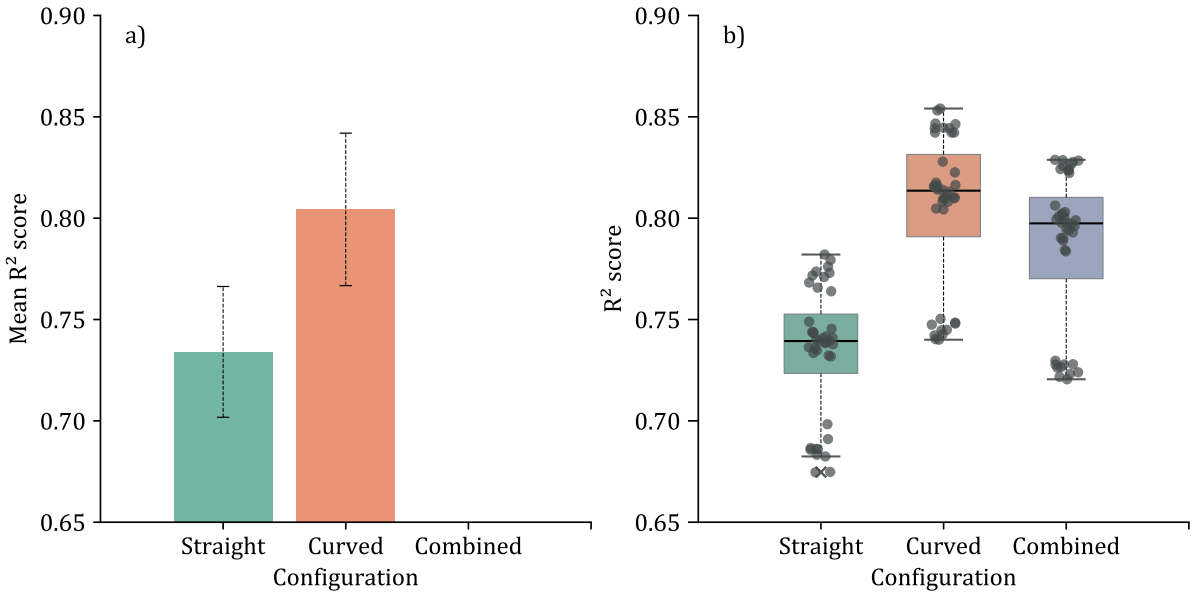


Figure 2.10: RF model validation performance across configurations using 10-fold cross-validation, illustrating the R^2 scores obtained for different configurations (straight, curved, combined). (a) displays the mean R^2 scores for each configuration, with error bars representing one standard deviation. (b) Combines a boxplot and swarmplot to visualise the distribution of individual R^2 scores across all folds. The boxplot highlights key statistical measures, including the interquartile range (IQR), whiskers that cover 95% of the data, and the median, indicated by a solid black line.

which the model captures more effectively. Additionally, the higher importance of G_{trail} in curved trails suggests that soil resistance plays a more significant role in non-linear machine routes. The relationship between vehicle dynamics and soil characteristics is more complex in curved configurations, resulting in stronger correlations between G_{trail} and rut development than in straight paths.

The RF model trained on observed rut depths was then applied to a synthetic spatial dataset composed of randomly distributed stumps and a simulated shear modulus map. The results showed that areas with smaller stumps generally had deeper predicted rut depths, while areas with larger and denser stumps showed shallower predicted depths (closer to 0 cm) across all three passes (Fig. 2.11a–c). This pattern also aligns with the model's feature importance results, where RRV and CI were identified as the most influential predictors (Fig. 2.9). Larger stumps likely indicate stronger, more extensive root systems, which help stabilise the soil and reduce deformation. Predicted rut depth also increased progressively with each machine pass, especially in areas with fewer stumps and lower shear strength. While the results are based on synthetic input data, they demonstrate how spatially explicit rut depth modelling can support forest operations planning. By identifying areas at greater risk of soil disturbance, managers can optimise machine routes, minimise long-term impacts, and support more sustainable harvesting practices.

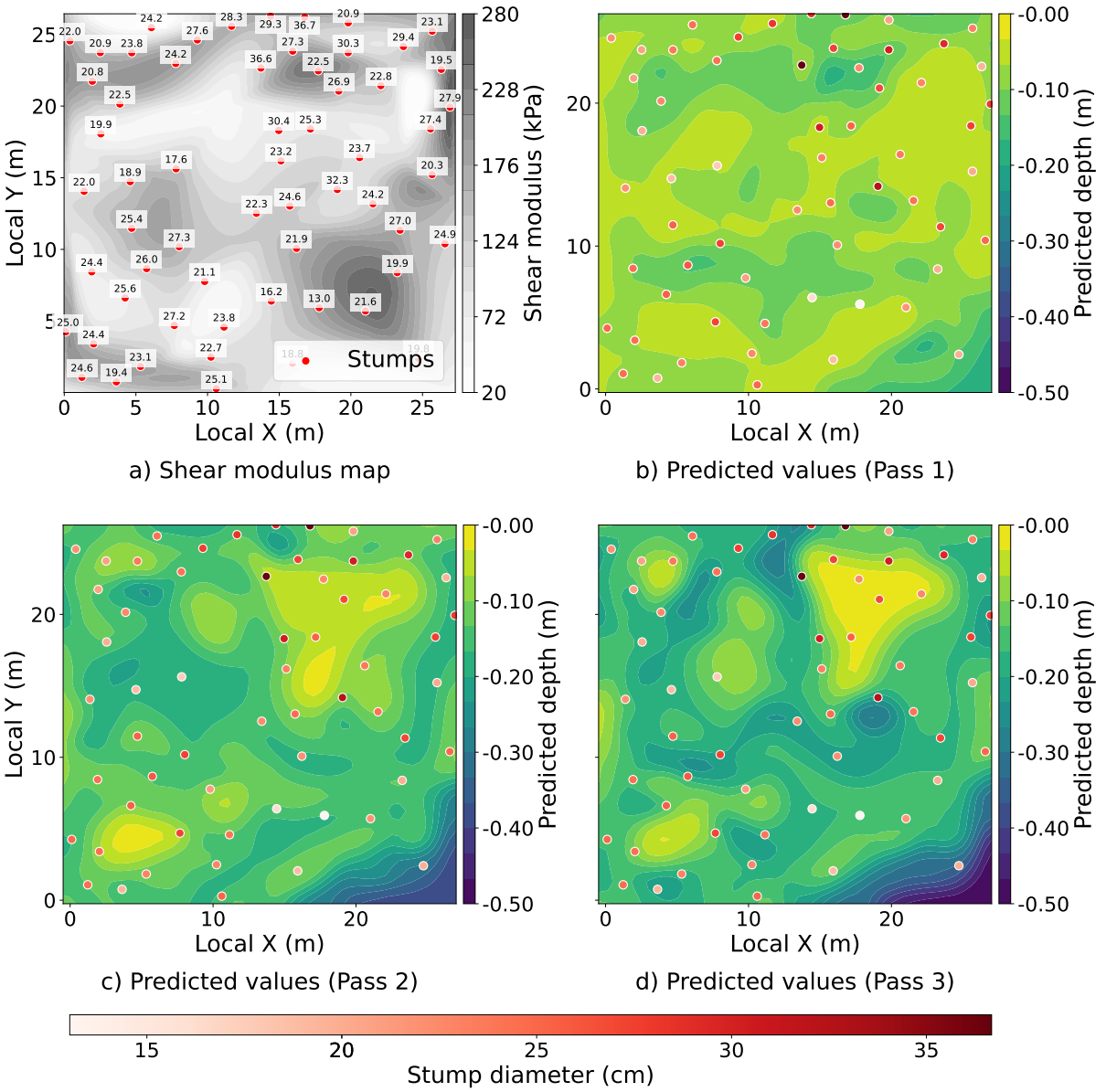


Figure 2.11: Spatial visualisation of shear modulus and predicted rut depth (cm) across multiple passes using a Random Forest (RF) model. Figure (a) presents a synthetic shear modulus map (in kPa) generated from randomly placed control points and interpolated across the area, overlaid with synthetic tree stumps (red dots). Figure (b–d) shows predicted rut depth values (for Passes 1, 2, and 3) using root reinforcement value (RRV), cumulative influence (CI), and shear modulus (G_{trail}) as predictors. The colour bar at the bottom represents the stump diameter (cm) and applies only to the red stump symbols. Each pass visualises rut depths alongside stump locations.

2.5 Discussion

This study investigated the role of tree stumps in rut formation on peatland soils, combining UAV-based and manual rut-depth measurements, statistical analysis, spatial modelling, and machine-learning techniques. The findings confirm that stump presence contributes to soil stability and mitigates machine-induced deformation. Rut depths were shallower near stumps, with the influence gradually weakening with distance. The impact of stumps was most significant in curved trails, where lateral soil displacement was greater, with ef-

fect sizes ranging from small to large (ε^2 = 0.02-0.32). Stump-related variables (CI and RRV) and *SM* were among the strongest predictors of rut depth. Machine learning models incorporating stump effects demonstrated strong predictive accuracy (R^2 = 0.69-0.85) in both straight and curved trails. These findings underscore the importance of considering stump presence in forest operations to improve machine path planning and support sustainable management. When used alongside DTW maps, these models could improve decision-making, minimise soil disturbance and reduce the environmental footprint of forestry activities.

The study first compared UAV-based and manual rut depth measurements to assess their correspondence. UAV-derived data provided consistent, high-resolution information and captured the same rut formation trends as manual measurements. Slightly lower values were observed in deeper ruts, particularly on curved trails and in later passes, with average differences of approximately 3 cm and maximum deviations of 3.3 cm. These differences likely reflect how each method interacts with the terrain more than inherent measurement inaccuracies. UAV-derived rut depths were extracted by averaging DTM values within a 10 cm area around each manual point, which may smooth out localised extremes. Moreover, photogrammetric models often define the ground surface at the top of low vegetation, such as moss and blueberry shrubs. At the same time, manual measurements are taken directly at the peat surface, potentially penetrating slightly into the soft, saturated peat, especially under wet conditions. This difference in reference level and surface interaction likely contributes to the observed discrepancies. Other contributing factors may include minor georeferencing offsets, interpolation limitations, and post-processing smoothing in photogrammetry software. In straight trails, where rut formation was more uniform, UAV and manual measurements agreed more closely. The higher RMSE (3.3 cm) and RMSE% (24.9%) observed in curved trails likely reflect a combination of these methodological factors and do not necessarily imply a systematic UAV bias.

Statistical analysis confirmed that cumulative stump influence significantly affected rut depth, with a small to medium effect in straight trails (ε^2 = 0.04-0.20) and a moderate to significant impact in curved trails (ε^2 = 0.02-0.32). The greater effect on curved trails is likely due to machines pressing down and pushing soil to the sides, causing greater soil displacement and deeper ruts. In these conditions, stumps can help stabilise the soil, possibly preventing excessive movement and reducing rut depth. In straight trails, the effect of stumps was smaller but still significant, as rut formation was primarily driven by vertical compression with less influence from lateral soil shifting.

This aligns with the feature importance scores and RFE analysis from the RF model, which identified stump-related variables (CI, RRV) and *SM* as the most influential predictors of rut depth. In contrast, NGP and peat strength did not significantly influence model performance. While these factors are generally crucial for rut formation, they were treated as site-specific constants, limiting their relevance to actual conditions. The varying ε^2 values could also be explained by the way CI and RRV were modelled. These variables assume

a uniform reduction in stump influence with distance, whereas rut depth typically exhibits a different pattern, being shallowest near the edges and deepest at the centre. When a stump is located to the side of a rut, CI and RRV decrease steadily, but rut depth first decreases toward the middle and then increases again. This inconsistency likely contributed to the lower effect sizes observed in some passes.

Another consideration is how the machine's weight distribution influences rut depth between the left and right wheel paths. Although left and right rut depths were measured separately, our model treated rutting as a combined process, modelled as a function of distance to the stumps. In practice, wheels may sink unevenly, especially when one side passes over a stump, leading to an asymmetric weight distribution and rut formation. When a stump lies directly in the wheel path, it acts as an obstacle the machine must climb over, often causing deeper ruts just before and after the stump due to increased resistance and slippage. Meanwhile, the opposite wheel is subjected to a higher wheel load, further contributing to the asymmetry in rut depth. Future studies could explore this by incorporating left- and right-rut depth averaging or a weight distribution model.

The RF models achieved high predictive accuracy (R^2 = 0.69-0.85), with better performance on curved trails than on straight trails. This was unexpected as straight trails were assumed to have more predictable rut patterns. However, the consistent rutting in straight trails made it harder for the model to detect subtle variations. In contrast, the greater variability in curved trails, likely due to lateral soil displacement, provided more diverse data for training, improving model performance. The model's application to the synthetic dataset showed that the predicted rut depth was generally shallower in areas with larger stumps and higher root reinforcement (Fig. 2.11). Rut depth increased with each simulated machine pass, especially in zones with fewer or smaller stumps. However, the pattern was not always consistent. Some areas with dense or larger stumps still showed deeper predicted rutting. This highlights the importance of how input parameters like RRV and CI are calculated and suggests that further refinement and customisation of these variables are needed. Nevertheless, these spatial predictions could support forest planning by identifying high-risk areas, optimising machine routing, and reducing long-term site disturbance.

Prior studies [73, 132] on UAV-based rut depth measurements have demonstrated the effectiveness of aerial photogrammetry in detecting soil deformations. However, they have also noted a slight underestimation of deeper ruts in complex terrain conditions [82]. Studies on machine-induced rutting emphasise the role of equipment design and soil properties in determining rut depth. Forwarder configuration, bogie tracks, and soil shear strength significantly affect rut formation, with heavier machines and wet conditions leading to deeper ruts [101, 105]. While these studies highlight the influence of machine and soil parameters [102], they often overlook stump-root systems as a natural mitigation factor.

The role of tree roots in soil stabilisation has been widely studied, with research demonstrating that root systems contribute significantly to mechanical reinforcement, particu-

larly on unstable or sloped terrain [114]. Similarly, studies on slope stability [112, 127] have demonstrated that tree roots increase shear strength and reduce lateral soil movement. This may explain why stumps exert a greater influence on rut depth in curved trails, where lateral soil displacement is greater.

While the stabilising effect of roots has been well documented, fewer studies have examined how stump-root systems affect soil deformation during forestry operations. [120] found that stump-root systems significantly reduced rut depth, with effects strongest within 0.5 to 1.5 meters from the stump. This supports the idea that stumps serve as natural reinforcements, with their impact gradually diminishing with distance [114]. However, previous studies have not incorporated spatial modelling to quantify these effects across larger areas, leaving a gap in understanding how stump influence varies across different terrains and machine traffic patterns.

While this study offers valuable insights, some limitations should be noted. UAV-based measurements, though effective, underestimated deeper ruts, particularly on curved trails, likely due to averaging within the 10 cm buffer and interpolation errors. Future research could enhance accuracy by incorporating LiDAR to capture fine-scale variations in rut depth.

The study's focus on a single peatland site may limit the generalizability of the results. Peat soils have unique properties, such as high organic content and water saturation, which may not be representative of mineral soils or drier forest conditions [94, 96, 98]. Expanding research to include diverse soil types and forest ecosystems would be necessary for broader applicability and to determine whether stump effects vary across ecological conditions.

Another limitation concerns the modelling of parameters (CI, RRV). In this study, we employed simplified spatial modelling techniques using UAV data and other parameters to quantify the effect of stump presence on rut formation. We assumed that reinforcement strength decreases gradually with increasing distance from the stump [114]. Although this offers a more straightforward way to represent root-soil interactions, it fails to accurately capture the spatial arrangement of root systems. Under real-world conditions, root expansion is inconsistent, varying by species, soil type, and tree age. Roots do not grow uniformly in every direction [115], suggesting that presuming a consistent, symmetrical decrease in reinforcement around the entire stump may lead to some level of overgeneralization. Also, calculating actual root reinforcement values is highly complex and typically requires extensive fieldwork and laboratory tests to measure parameters such as root tensile strength, root area ratio (RAR), soil shear strength, and root distribution [114, 119, 133]. Although our method establishes a useful experimental baseline, it may not fully capture the unique root architectures across species or the mechanical characteristics of specific root systems. Future studies might improve model precision by developing species-specific constants for root reinforcement assessments, particularly for economically significant species like Scots pine and Norway spruce. This could be accomplished by integrating field-based root

mapping, which could help create species-specific root spread patterns. Techniques such as ground-penetrating radar (GPR) have already been successfully employed in ecological and geotechnical studies to identify and visualise three-dimensionally shallow and deep tree root systems [134].

The stabilising influence of stumps on soil support, particularly in reducing rut formation, highlights their critical role in sustainable forest operations. By incorporating stump-related variables into predictive models, forest managers can improve rut depth forecasting, optimise operational planning, minimise soil disturbances, and improve both economic and ecological outcomes. These models can help identify optimal skid trails and areas to avoid, reducing unnecessary machine movement and lowering the risk of damage to machinery. This approach could also improve planning efficiency and reduce operational costs. However, as the current approach to root reinforcement modelling remains simplified, improvements are necessary. These advancements will contribute to more sustainable forest management practices that balance operational efficiency with ecological preservation.

2.6 Conclusion

This study demonstrated that stump presence plays an essential role in reducing rut formation on peatland soils, with effects varying by trail configuration. Using UAV-based measurements, statistical analysis, spatial modelling, and machine learning, we highlight the role of biological reinforcement in mitigating soil deformation induced by machines. The RF model identified stump presence, reinforcement effects, and shear modulus as key predictors of rut depth, demonstrating the combined impact of biological and mechanical soil properties.

Integrating stump influence into rut depth prediction models can improve machine path planning and soil conservation in peatland forestry. However, site-specific conditions and simplified reinforcement modelling suggest that species-specific constants and advanced mapping techniques, such as GPR, could improve accuracy. Future research should refine root reinforcement models, integrate high-resolution root mapping, and expand to diverse soil types to improve predictive accuracy and support sustainable forest management.

2.7 Data availability

The datasets generated and analysed in this study are available in the University of Padova Research Data repository at `https://researchdata.cab.unipd.it/1722/`.

The corresponding published paper is available at `https://www.sciencedirect.com/science/article/pii/S1537511025001916`.

CHAPTER 3

TERRAIN ROUGHNESS MAPPING USING UAV PHOTOGRAMMETRY

3.1 ABSTRACT

In forestry, assessing and navigating the terrain is crucial for operational efficiency. Terrain trafficability, the capacity to support vehicle passage, affects machinery movement and soil health. While soil's weight-bearing ability is crucial, terrain classification is equally essential for defining an area's characteristics. This study automates terrain-roughness estimation from UAV imagery to generate high-resolution roughness maps. We used deep learning for object segmentation, followed by point-cloud classification and ground-roughness quantification. Manual field measurements validated the segmentation results, while manipulating the obstacle dataset tested the sensitivity of the roughness algorithm. In addition, drone-derived digital terrain models (DTMs) were used to calculate the terrain ruggedness index (TRI), vector ruggedness measure (VRM), and area ratio (AR) for comparison. Obstacle segmentation achieved 95.6% accuracy, while height estimation had an RMSE of 2.6 cm and an MRE of 11.4%. Manipulation of the data set demonstrated the method's responsiveness to changes in obstacle density and height. The trends in TRI and AR values ($rho = 0.63, p < 0.05$) and ($rho = 0.67, p < 0.05$) indicate that the method classifies areas similarly to TRI and AR. In contrast, VRM (ρ = 0.24, ρ = 0.13) did not align well with the roughness method. This study highlights the potential to automate and improve roughness assessment, thus enhancing operational efficiency and allowing more accurate adjustments to performance expectations and cost estimates in forest operations.

3.2 INTRODUCTION

Heavy machinery, such as harvesters and forwarders, is essential in forestry, and their performance depends heavily on soil bearing capacity [135]. Trafficability, defined as the soil's ability to support machinery [2], is critical for ensuring efficient operations. In recent decades, the widespread adoption of powerful, large-scale forestry machinery for logging operations and transportation in industrial wood procurement [35, 136] has high-

lighted the growing need for a deeper understanding of terrain trafficability. Trafficability is often categorised by soil type, moisture conditions, surface characteristics, and seasonal variations [49, 137]. Poor trafficability can lead to delays and increased operational costs, and cause environmental damage, such as soil compaction and rutting, which affect tree growth and disrupt water management [47].

Beyond the soil's ability to support machinery, trafficability also includes navigability, the ease with which vehicles can traverse the terrain. Factors such as surface roughness and physical obstacles play a critical role in vehicle performance [11, 138]. To address these challenges, TCS have been developed to evaluate and categorise forest land accessibility, enabling more efficient forest operations [1]. TCS is a tool that helps forest managers evaluate terrain characteristics, plan budgets, control costs, and allocate machinery to the most suitable locations for forest operations. This classification depends on three key factors: a) ground conditions, which include the soil's bearing capacity, which is influenced by soil type and moisture; b) ground roughness, referring to obstacles such as rocks and depressions; and c) slope condition, classified according to the gradient and topographic form of the terrain [1]. These factors provide valuable information on the terrain and support effective operational planning. For example, assessing ground conditions can help determine whether the soil can support heavy machinery without causing damage, while evaluating slope conditions ensures that steep areas are navigable or avoided for safety.

In recent decades, several TCSs have been developed to systematically assess these factors [1, 41, 44, 40, 43]. These approaches are generally categorised into two main types: descriptive and functional classifications. Descriptive classifications group sites based on physical features such as soil moisture, bearing capacity, slope, and roughness. This approach focuses on classifying the terrain based on the characteristics of the permanent site rather than the type of machinery used. By relying on natural features, it remains adaptable to new harvesting equipment and eliminates the need for repeated evaluations [40].

In contrast, functional classifications evaluate the terrain's suitability for specific types of machinery, directly addressing the practical needs of forest operations [42]. Traditionally, these assessments have been performed manually, which requires expertise and experience to evaluate factors such as slope, ground roughness, and soil conditions on large plots of at least 100 square metres [1]. Although effective, this method faces several challenges, including inconsistencies in observer skill, subjective judgment, and the time required to collect the data [1, 40]. In addition, the manual nature of these assessments makes them inefficient for large-scale forest operations and can delay crucial decision-making.

Terrain roughness has been the focus not only in forestry but also in various other disciplines, where its definition and application often vary across contexts. In disciplines such as geomorphology, road surface assessment, and ecology, roughness is typically described as a measure of surface complexity, quantifying elevation variability or surface uneven-

ness [139, 140, 141, 142]. Metrics such as the vector ruggedness measure (VRM) [143], terrain ruggedness index (TRI) [144], and local elevation deviations are frequently used to quantify roughness in studies of landslides [141], natural hazards, and landform classification [79]. Typically, these metrics are derived from digital elevation models (DEMs) and provide a generalised understanding of terrain characteristics.

In forestry and rural landscapes, terrain roughness is critical for road conditions and wheel rutting, which is caused by vehicle traffic on soils with low bearing capacity or high moisture content. Rutting can alter hydrology, compact the soil, and hinder vegetation growth [145]. Furthermore, road surface roughness can significantly influence vehicle performance and fuel consumption, with studies showing a 107% increase in fuel consumption between the highest and lowest road quality classes [146]. Terrain roughness is also an important indicator of ecological complexity. Areas with high roughness are often associated with increased species richness and biodiversity hotspots, reflecting the diverse habitats and environmental gradients created by varied topography [147].

Advances in technology, such as UAVs (unmanned aerial vehicles) and LiDAR (light detection and ranging), have enabled more efficient and scalable methods for assessing terrain characteristics [46, 47]. In the past decade, UAVs have become increasingly common for assessing soil conditions and roughness, providing cost-effective alternatives to traditional ground-based methods [52]. UAV photogrammetry has been used effectively to quantify soil surface roughness from UAV-derived DEMs [48]. Additionally, UAV datasets have been used to compute TRI and VRM at flight altitudes ranging from 20 to 360 metres, indicating that higher-resolution digital terrain models (DTMs) retain more detailed surface-roughness information [148]. Furthermore, UAVs have supported multi-scale analyses of terrain roughness to improve natural hazard modelling. For example, DSMs derived from UAV photogrammetry and LiDAR data in the European Alps have accurately described terrain characteristics such as slope gradients, surface roughness, vegetation cover, and channelised flow paths, all of which are critical for hazard prediction [79]. Despite the effectiveness of UAVs in many applications, they often struggle to capture detailed surface information, whereas LiDAR technology offers a distinct advantage [26, 49]. LiDAR, including ALS (airborne laser scanning) and TLS (terrestrial laser scanning), has been effectively used to assess and predict terrain conditions. ALS has been utilised in Finland to evaluate terrain trafficability, thereby identifying areas at risk of soil compaction from vehicle traffic [49]. In another study, Vega et al. [26] employed ALS to collect elevation data in northwestern Spain and developed a modular terrain model to assess daily variations in machine-specific forest soil trafficability. Wallin et at. [149] utilised high-resolution airborne LiDAR data from the SCA Laxsjö digital test site in Sweden to develop a model capable of predicting terrain traversability with 90% precision. Their work focused on three critical performance metrics - target speed, energy consumption, and acceleration, key factors in vehicle efficiency and environmental impact.

TLS, on the other hand, has been used for more localised applications, such as eval-

uating soil surface roughness by capturing high-resolution, two-dimensional data and analysing the influence of scan positions on the roughness indices [150]. It has also been used in geohazard mapping, including slope stability assessments and landslide potential analysis [151]. Dynamic maps integrating hydrological models and sensor data have been applied in Finnish forests to predict trafficability and mitigate environmental impacts, such as soil rutting [56]. Another approach involves using GIS-based models and depth-to-water maps to improve trafficability assessments [47, 152]. However, a crucial factor to consider is that the high costs of TLS and LiDAR technology, in general, can limit their use in smaller-scale studies or budget-constrained projects.

Although these methods are effective in many cases, deep learning can improve these processes by automating tasks and improving object identification. A key aspect of deep learning is segmentation, which enables the extraction of more detailed and accurate information from images. In segmentation, each pixel in an image is assigned to a specific region based on the algorithm's parameters, thereby simplifying the scene's complexity. Recent advances in deep learning models have achieved high accuracy in analysing visual content. For example, models like SAM (SegmentAnything) achieve high accuracy and exhibit zero-shot generalisation to unfamiliar objects and images. SAM uses the bision transformer (ViT) architecture, originally designed for image classification. ViT processes images in patches and employs self-attention mechanisms to capture long-range relationships.

Additionally, ViT has shown strong performance in semantic segmentation tasks, enabling visual elements to be grouped according to segmentation parameters without relying on predefined classes [153, 154, 155]. However, despite these advances, the application of such techniques to automate terrain classification within the TCS framework remains limited. Furthermore, while terrain roughness has been widely studied, research has often been application-specific, with fewer studies focusing specifically on roughness within TCS, particularly regarding ground conditions, slope, and detailed terrain characteristics outlined in the national TCS [1].

Recognising the importance of such data, our study aims to automate terrain roughness assessment using UAV imagery, specifically in a post-fire area of the Jonkershoek timber plantation, South Africa. By leveraging the unique surface conditions of the experimental site, we aim to develop and validate a method for generating high-resolution roughness maps at a granular 3x3-metre scale. The primary objective is to demonstrate how UAV-derived data can streamline and improve traditional, labour-intensive terrain classification approaches, which currently rely on manual evaluation. Although we acknowledge some limitations, we also outline plans to improve the method's broader applicability.

3.3 MATERIALS AND METHODS

This section describes the materials and methods used to derive grid-based terrain roughness classes from UAV RGB imagery, including field data collection, photogrammetric re-

construction, obstacle detection, roughness classification, and statistical comparison with established terrain metrics. An overview of the general processing workflow is provided in Fig. 3.1.

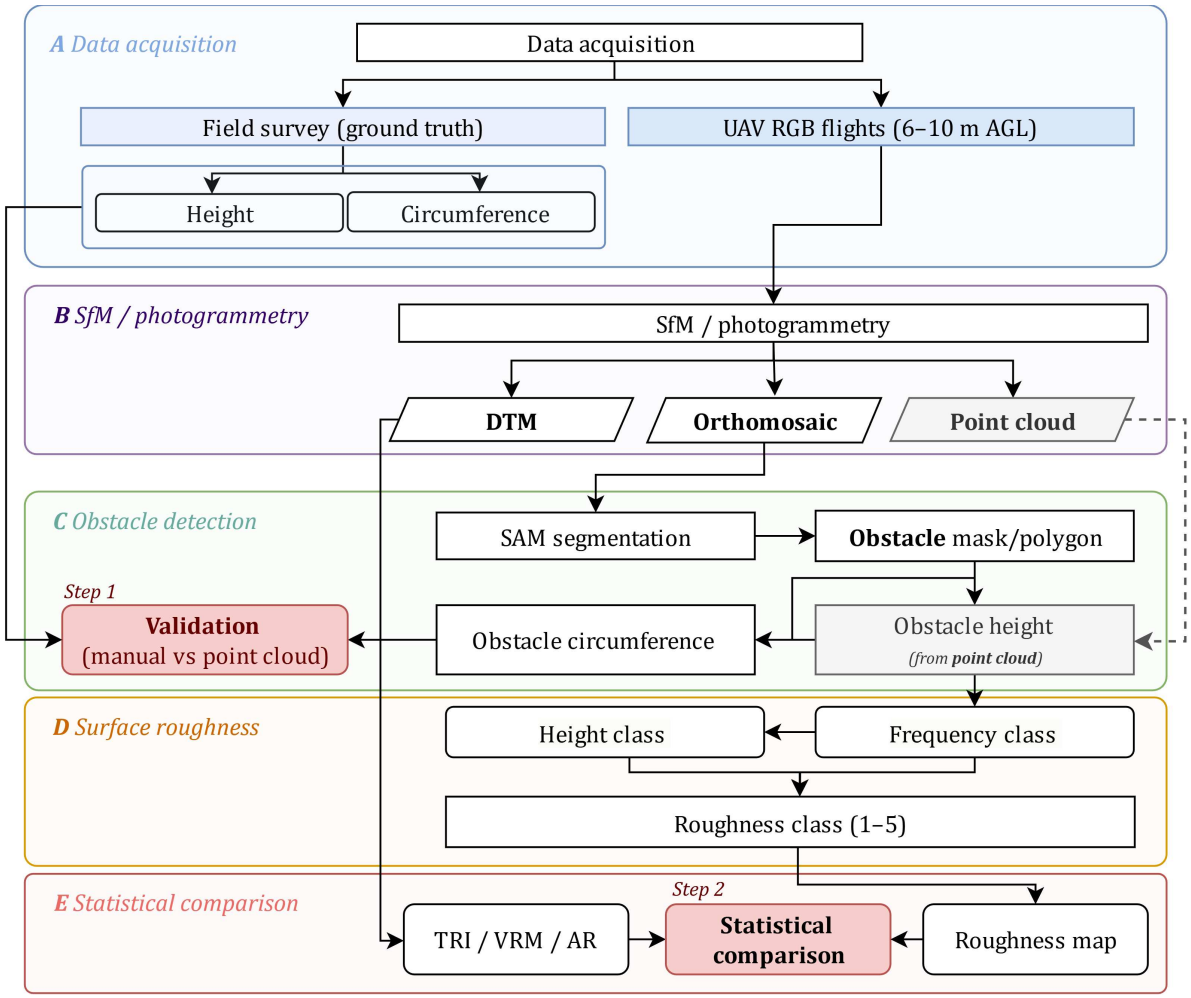


Figure 3.1: Workflow used to derive grid-based roughness classes from UAV RGB imagery. SfM outputs (orthomosaic, DTM, and point cloud) support SAM-based obstacle segmentation, with the obstacle circumference derived from polygons and the obstacle height extracted from the normalised point cloud. Field measurements are used to validate obstacle metrics, and resulting roughness classes are compared with DTM-derived indices (TRI, VRM, AR).

3.3.1 STUDY AREA

The research was conducted in a post-fire area within the Jonkershoek timber plantation in the Western Cape province of South Africa (33°58'19.61"N, 18°56'42.24"W; UTM32S (Fig. 3.2). The region experiences a Mediterranean climate characterised by warm, dry summers and cool, wet winters [156]. These climatic conditions, along with elevated summer temperatures and regular south-easterly winds during summer and fall, create a high fire risk. In addition, warm and dry winds in spring further increase the likelihood of fire outbreaks [157, 158].

The chosen study area was affected by a forest fire that significantly altered vegetation and terrain characteristics. It is now characterised by sparse or no vegetation and a rocky surface with rocks of varying sizes. Conventionally, the site would have been harvested motor-manually, with tree lengths extracted by a cable skidder. Using this controlled post-fire setting as a baseline, our objective is to also adapt these applied methods for future studies in more complex environments, such as areas with partial or full canopy closure.

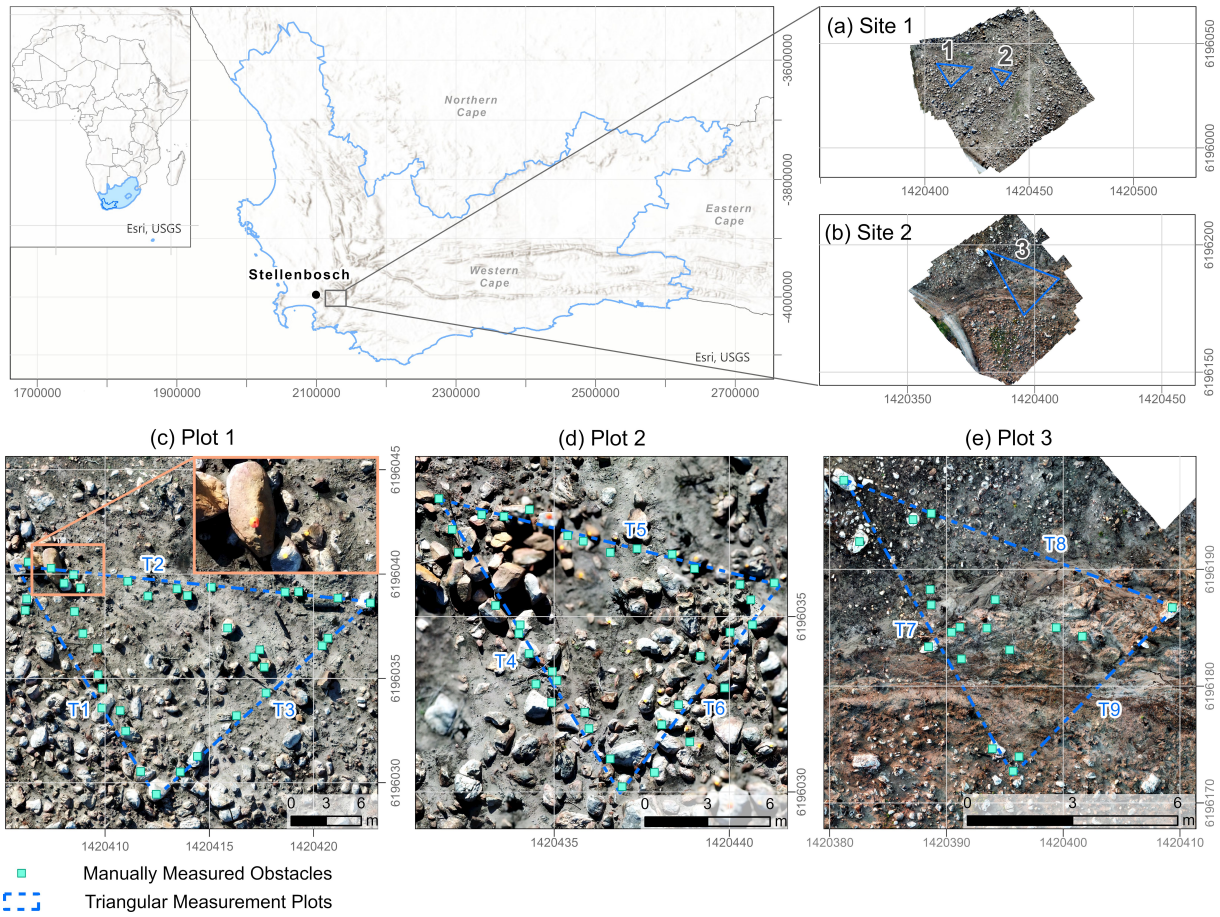


Figure 3.2: Study site and plots along with measurement methodology. (a, b) The study sites with triangular plots are located in the Western Cape province of South Africa. (c-e) Detailed views of the measurement plots show manually measured obstacles (green squares) and triangular measurement areas (blue dashed lines). Obstacle measurements were systematically recorded along three transects forming triangles at the centre of each plot: plot 1 (T1, T2, T3), plot 2 (T4, T5, T6), and plot 3 (T7, T8, T9). Site 1 is later used for roughness estimation.

3.3.2 FIELD AND UAV DATA COLLECTION

A single field campaign was conducted to collect ground-truth field data from two study sites, covering approximately 0.75 hectares. The first study site, which covered 0.5 hectares, contained two triangular plots (Fig. 3.2a), while the second site, measuring 0.25 hectares, consisted of one triangular plot (Fig. 3.2b). Across the three triangular plots, 90 obstacles (rocks) were selected and marked for identification in drone imagery (Fig. 3.2).

The obstacles were marked with spray paint to enhance their visibility in the drone orthophotos. Their locations followed a sequential pattern along predefined transect lines, starting at the triangle's corner and continuing along its sides. This arrangement enabled accurate reconciliation of their measurements during analysis.

The selected obstacles were measured in circumference and height using a field measuring tape. Measurements were systematically recorded along these predefined transect lines within each study plot. The transects varied in length between plots, averaging 13.5 m in plot 1 (T1, T2, T3: 15.2 m, 13.6 m and 11.8 m; Fig. 3.2c), 8.9 m in plot 2 (T4, T5, T6: 9.7 m, 7.3 m, and 9.7 m; Fig. 3.2d), and 14.5 m in plot 3 (T7, T8, T9: 13.6 m, 12.9 m, and 17.2 m; Fig. 3.2e). The triangular layout was designed to capture variations in obstacle size and distribution across the study area, ensuring a representative sample of the rocky environment. Most obstacles along designated transects were measured, and each was assigned a unique identifier. This included a 'transect ID' to identify the measurement line and an 'obstacle ID' to reference each obstacle. Table 3.1 summarises the obstacle measurements within the surveyed plots.

Table 3.1: Summary of obstacle measurements by site

Plot	Number of measured obstacles	Circumference (cm)			Height (cm)		
		Min	Max	Mean	Min	Max	Mean
Plot 1	36	50	410	179	5	75	27
Plot 2	35	60	315	167	4	68	25
Plot 3	19	72	770	243	11	74	30
Total	90	50	770	187	4	75	27

The primary objective of this study was to automate the estimation of ground roughness and to define a corresponding methodology. Manual field measurements collected during the field campaign primarily served as benchmarks for validating a deep learning segmentation algorithm, which is crucial for obstacle detection. These data help evaluate and improve the algorithm's accuracy in surface-roughness estimation, thereby reducing reliance on traditional, labour-intensive methods.

During the same field survey, three drone flights were conducted under clear skies - one for each triangular plot - to capture high-resolution images of the study area. A DJI Mavic Pro drone equipped with a DJI FC220 camera was used to collect data. To achieve fine resolution, the drone was operated at low altitudes of 6 to 10 metres above ground level, maintaining an image front overlap of 80% and a side overlap of 70% for complete coverage and data accuracy. In total, 90, 78, and 45 images were captured for plots 1, 2, and 3, respectively.

The collected images were processed to generate orthomosaics and dense point clouds. Since plots 1 and 2 were located close together (Fig. 3.2a), they were processed as a single unit to cover a larger area. In contrast, plot 3, being farther away, was processed independently (Fig. 3.2b). Image preprocessing was performed using Agisoft Metashape Professional [122], which employs Structure from Motion (SfM) and photogrammetric stereo-matching techniques to reconstruct 3D models from unordered but overlapping images.

For spatial alignment, we used a straightforward approach based on global positioning system (GPS) information embedded in the RGB image metadata, without integrating ground control points (GCPs). Instead, the data were processed in a local coordinate system. After image alignment, three fixed corner points of each triangular plot were manually assigned as markers, and field-measured distances between them were used to apply a scale constraint in the XY dimensions. Since there was no altitudinal difference among these points, the Z dimension was assumed to be consistent within the study area. This ensured that the objects in the orthomosaic were correctly scaled relative to their actual size.

Although this approach provided sufficient data for method development, it comes with trade-offs, particularly regarding absolute elevation accuracy and potential systematic errors or shifts in the final orthomosaic or DTM. To assess the reliability of the Z dimension and achieve the highest possible precision, we conducted a preliminary test in a small area within the study site to obtain the most precise height data. We used 67 manually measured obstacle heights as benchmarks and compared them with those derived from the drone-generated point cloud under different processing settings. This comparison enabled us to evaluate the effects of various quality settings and depth-filtering techniques in Agisoft Metashape Professional on the results. We tested all combinations of quality settings (high, medium, and low) and depth filtering modes (aggressive, moderate, and mild) [122].

To quantify these differences, we calculated the mean relative error (MRE), SD, and RMSE (Eqs. 3.1–3.3) for the manual and predicted heights across multiple configurations (e.g., aggressive depth filtering and medium-dense point-cloud generation settings).

$$MRE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{P_i - T_i}{T_i}\right| \tag{3.1}$$

$$SD = \sqrt{\frac{\sum(P_i - T_i - \bar{D})^2}{n}} \tag{3.2}$$

$$RMSE = \sqrt{\frac{\sum(P_i - T_i)^2}{n}} \tag{3.3}$$

where $\bar{D}$ is the mean of the differences between the predicted heights P_i and the actual heights T_i. P_i represents the predicted heights from the drone, T_i represents the actual

heights from manual measurements, and n is the number of measurements.

Point clouds were exported in standard .las format, and both a DTM and an orthomosaic image were produced with a very high resolution of 0.05 metres, determined by the average ground sampling resolution of the source photos. The final products, the DTM and the orthomosaic, were exported as .tif files. Consequently, dense point cloud data were utilised to extract the heights of the segmented obstacles.

All roughness estimation processes and analyses were coded in Python, integrating tools such as LAStools [159] and FUSION/LDV [160], and were executed automatically within the Python environment. UAV image processing and point cloud generation were performed using Agisoft Metashape Professional software [122].

3.3.3 Obstacle segmentation

While traditional remote sensing segmentation methods rely on pixel- or object-based approaches, they often struggle with the complexity of modern applications [161]. In this study, we focused on achieving high-quality segmentation of complex surface structures. We employed the SAM (Segment Anything) deep learning framework for its simplicity, computational efficiency, and competence in detailed instance segmentation [155]. We used SAM to segment and extract obstacle polygons, which were subsequently used for point cloud classification and surface-roughness estimation. Several preprocessing steps were performed to ensure accurate segmentation. Orthomosaic was segmented into 2560-pixel tiles with 20% overlap (512 pixels).

Additionally, the overlap helps prevent gaps in feature recognition across the dataset. We implemented the segment-geospatial (samgeo) framework using the ViT_H (vision transformer-huge) model. This model processes images as sequences of patches, enabling it to capture global dependencies and details across the image through self-attention mechanisms. Known for its ability to identify fine-grained details, the ViT_H model effectively detects subtle features in high-resolution imagery [153, 154]. We further fine-tuned the SAM model to adapt it to our specific dataset. This process involves adapting a pre-trained model to new data, thereby improving its ability to handle the unique characteristics of our dataset. By manually annotating 145 images in our custom dataset using the LabelMe tool [162], we created 458 initial segmentation masks. Each image mask file was encoded as a binary image: pixels with value 0 indicate no object, and those with value 255 indicate object presence. After deriving bounding boxes from the segmentation masks, we implemented a class to manage this dataset. This class simplifies handling input images and masks, pairing them for further processing. Following this setup, we trained our model and periodically saved its state. To achieve optimal results, we employed a systematic trial-and-error approach to refine the model's parameters, thereby enhancing its predictive performance on our specific dataset. We then applied the trained model to raster tiles for segmentation, converting the results into shapefiles for further analysis across the study area. Each shapefile contained separate polygons for each identified feature. To ensure

high-quality segmentation results, incomplete polygons at tile borders in the overlapping zones between adjacent tiles were automatically removed using a custom Python script.

Additionally, redundant features, such as small polygons or elongated polygons representing burned logs, were automatically removed from the entire obstacle dataset. This approach allowed us to preserve all relevant obstacles while removing incomplete or irrelevant features. Due to the complex terrain, the segmentation algorithm occasionally captured features other than the target, such as low vegetation. We applied an RGB colour threshold to refine our results and exclude areas with green foliage. This threshold was defined by manually selecting 100 polygons of green vegetation and extracting their RGB values to represent a range of RGB combinations. This range was then applied to the dataset to filter out vegetation, and the resulting filtered dataset was manually reviewed to ensure accuracy.

3.3.4 Point cloud classification

We used polygons detected by the model to clip point clouds into ground and non-ground points using the FUSION/LDV *PolyClipData* function, ensuring the separation of relevant terrain features. To further refine this classification and address missed obstacles, we applied LAStools *lasground* with the "*-nature*" and "*-fine*" parameters, optimised through trial and error and visual inspection. We employed *las2las* to filter newly classified ground points, which were merged with the original ground points using *lasmerge*, resulting in a more accurate and refined ground point cloud. This enhanced ground point cloud was classified as Class 2 using *lasclassify*. Non-ground points were merged with the original obstacle polygon-clipped point cloud and classified as Class 0. Both ground and non-ground point clouds were then merged for comprehensive analysis. Finally, we used *lasheight* to remove noise below -0.1 meters and above 10 meters and to normalise the point cloud for heights above the ground surface. The *lasheight* tool normalises the point cloud by subtracting the digital terrain model (DTM) elevation from each point's elevation. Tiling with LAStools *lastile* ensured manageable data sizes throughout the process.

3.3.5 Obstacle-based roughness mapping

3.3.5.1 Terrain classification and roughness assessment

Traditionally, roughness assessment is conducted visually on larger plots of at least 100 square meters [1]. To improve precision, we refined the terrain classification method by applying it to smaller 3 × 3-meter cells, enabling more detailed assessments that can later be scaled up or down. In this case, the 3 × 3-meter grid provides a practical starting point for assessing terrain roughness, identifying obstacles, and computational efficiency. It approximates navigation conditions for forest machinery but does not necessarily represent the minimum space required for manoeuvrability. The classification outcome is a grid-based roughness map, in which each 3 × 3-meter cell is assigned a roughness value.

The terrain analysis involves two key metrics: obstacle height, derived from the *.las* point cloud data, and obstacle frequency, determined from the spatial distribution of obstacles represented as polygons. Stumps and logging residue, which will degrade over time, are not counted as obstacles in the context of a descriptive terrain classification.

Within each 3 × 3-meter grid cell, every obstacle is classified according to established rules. Fig. 3.3 visually represents this process. The logic evaluates each feature based on its height and proximity to nearby obstacles, determining whether it is a significant or insignificant obstacle, according to specific thresholds and spatial rules.

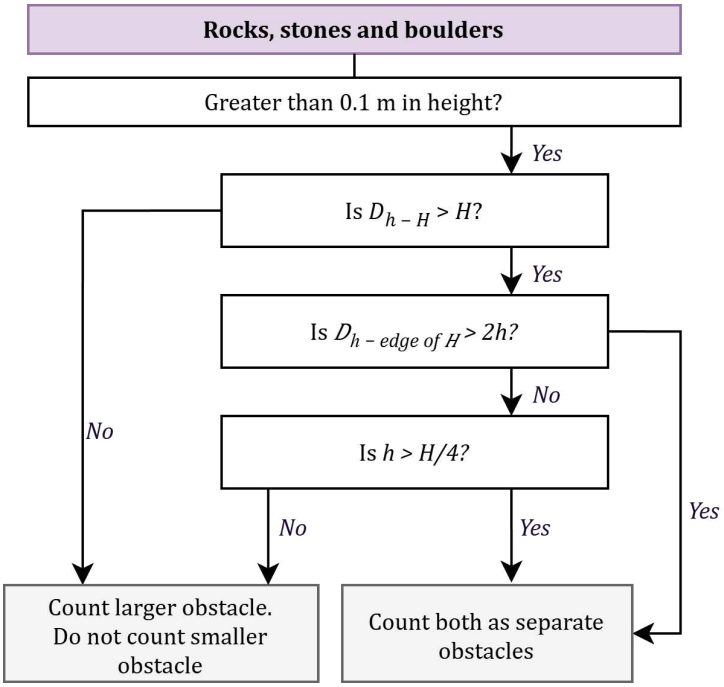


Figure 3.3: Flowchart for obstacle counting based on height and proximity. The diagram presents the decision-making process for classifying obstacles (rocks, stones, and boulders), where h is the height of a small obstacle, H is the height of a large obstacle, s is a small obstacle, L is a large obstacle, D is distance, D_{s-L} is the distance from the center of the small obstacle to the center of the large obstacle, and $D_{s \text{ to edge of } L}$ is the distance from the center of the small obstacle to the edge of the large obstacle.

3.3.5.2 Obstacle counting

The classification begins by assessing each obstacle's height and assigning it to a height class according to predefined limits. Obstacles are categorized into four groups: H20 (0.10–0.29 m), H40 (0.30–0.49 m), H60 (0.50–0.69 m), and H80+ ($\geq$ 0.70 m) [1]. Obstacles less than 0.1 meters are considered insignificant and excluded from further analysis. Initially, the dataset contained 8936 segmented obstacles. Applying the height-based filtering criteria reduced the number of obstacles to 4246 for further analysis. The overall statistics of the obstacle dataset showed that the average obstacle height is approximately 31 cm, with a range of 10-120 cm (Table 3.2).

For those exceeding the 0.1-meter threshold, a proximity buffer equivalent to their height was created around each obstacle's centroid. This buffer helps check for nearby obstacles. Any obstacle without neighbours in this buffer is counted as a separate obstacle.

Table 3.2: Summary of obstacle height statistics for the original dataset, categorised by height class (H20, H40, H60, H80+).

Height class	Count	Mean (cm)	Min (cm)	Max (cm)
Total	4254	31	10	120
H20	3169	16	10	30
H40	722	38	30	50
H60	263	59	50	70
H80+	92	89	70	121

For each smaller obstacle within the proximity buffer of a larger one, we measure the distance from the smaller obstacle's centre to the edge of the larger obstacle's buffer. If this distance exceeds twice the height of the smaller obstacle, it is counted separately for surface roughness. If not, we compare their heights: if the smaller obstacle's height is less than a quarter of the larger one, it is not counted separately to prevent skewing the roughness assessment. Otherwise, both are counted independently. Finally, obstacles are categorised into frequency classes based on height and average distances, calculated using Delaunay triangulation.

3.3.5.3 Obstacle classification

Due to the dimensions of the 3 × 3-meter grid cells, we refined the original frequency estimation system, which was initially designed for larger areas and calculated per hectare. Applying these hectare-based thresholds directly to a 3 × 3-meter grid is impractical, so we derived new thresholds proportional to the smaller grid size.

As described by Erasmus [1], the original classification system defines four frequency classes based on the distance between obstacles and the number of obstacles per hectare. Obstacles are considered isolated when the distance between them is > 16.0 m, with < 40 obstacles per hectare. The infrequent class applies to distances of > 5.0 m and ≤ 16.0 m, with 40 to < 400 obstacles per hectare. Obstacles are moderately frequent when distances range from ≥ 2.2 m to ≤ 5.0 m, at densities of 400 to 2000 obstacles per hectare. Obstacles are considered frequent when the distance between them is < 2.2 m, with > 2000 obstacles per hectare.

Due to its practicality, a 3 × 3-meter grid was chosen as the basis for refining these frequency classes, with a 9 m^2 area. The scale factor was extracted by dividing the area of a hectare by the area of the 3 × 3 meter grid (Eq. 3.4).

$$\mathrm{SF} = \frac{A_{\mathrm{unit}}}{l^2} \tag{3.4}$$

where A_{unit} is the area of the standard unit (e.g., $10,000\ \mathrm{m}^2$ for a hectare) and l is the side length of the grid in meters. Using this SF, we reduced the number of obstacles per hectare to a 9 m^2 area (Table 3.5).

We used the lower quantile (Q1), median (Q2), upper quantile (Q3), mean, and SD from the rescaled values to establish obstacle frequency thresholds. Q1 and Q2 capture the lower tail of the distribution, ensuring that lower-frequency classes are well-defined. Q3, combined with the mean and SD, captures the upper end, accounting for higher variability and defining moderately frequent and frequent classes (Table 3.5). This ensures that the obstacle density remains consistent with the original classification and applies to smaller areas.

To accurately identify cells with large obstacles and avoid misclassification, we assigned weights to each grid cell based on obstacle height and the area they occupied within each cell (Eq. 3.5). The weights were defined as follows: H20 (1.0), H40 (1.2), H60 (1.5), and H80+ (2.0) and calculated using the following formula:

$$W_{hc} = OC_{hc} \times w_{hc} \times A_{hc} \tag{3.5}$$

where W_{hc} is the weighted obstacle count for height class hc, OC_{hc} is the obstacle count for height class hc, w_{hc} is the weight assigned to height class hc, and A_{hc} is the area occupied by obstacles of height class hc, in m^2.

Table 3.3: Adjusted obstacle incidence classification for a grid size of 3×3 meters.

Category	Original rescaled value (per 9 m^2)	Calculation	Adjusted value
Isolated	0.036	Q1 + mean	0–1
Infrequent	0.36	Q2 + mean	2–3
Moderately frequent	1.8	Q3 + mean	4–6
Frequent	2.7	Q3 + std	>7

To address edge effects, we employed a shifted-window approach by shifting the grid by half the cell size in the positive and negative X (right and left) and Y (up and down) directions to create overlapping regions (Fig. 3.4). This method allows handling obstacles on the boundaries of grid cells. We calculated the frequencies for each cell in the original and shifted grids. We then overlaid the grids to determine the final frequency for each 3 × 3 m cell. A rule was applied to determine the frequency value (the most common value, or the highest value if multiple values are equally common).

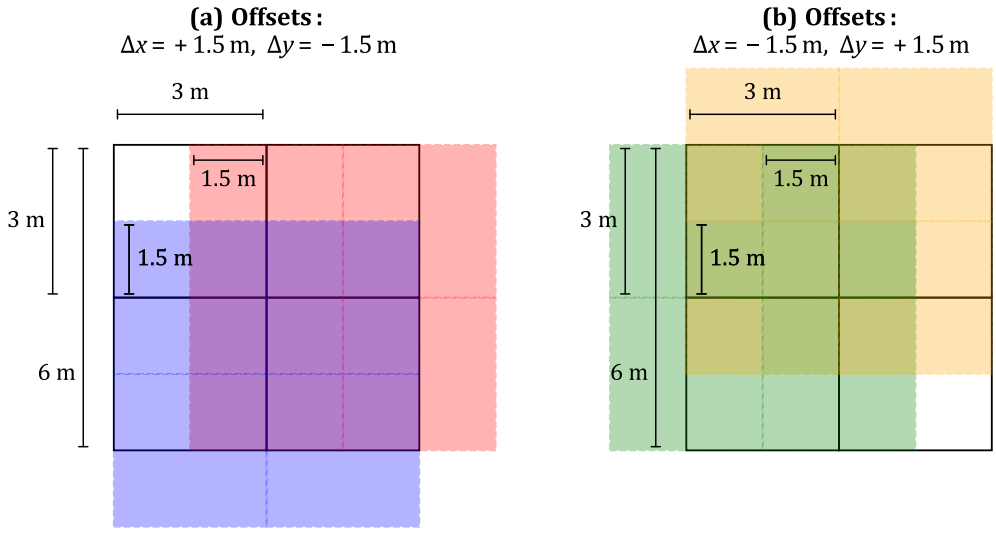


Figure 3.4: Shifted-window approach to reduce obstacle edge effects. (a) Offsets of $(\Delta x, \Delta y) = (+1.5\,\text{m}, -1.5\,\text{m})$ and (b) $(\Delta x, \Delta y) = (-1.5\,\text{m}, +1.5\,\text{m})$. Base window size is 3×3 m.

3.3.5.4 ROUGHNESS CALCULATION

To classify ground roughness within each cell, we use a tabulated method that integrates obstacle-frequency and obstacle-height classification. Starting with the smallest height class, the process moves horizontally across a structured table, adjusting from H20 to H80+ (Table 3.4). This systematic approach ensures all relevant data are considered, providing a quantitative classification from "*smooth*" to "*very rough*" and avoiding the subjectivity of visual evaluations.

To test the algorithm's sensitivity, we conducted a series of tests by systematically manipulating the obstacle dataset. We aimed to understand how the algorithm responds to changes in the dataset. To simulate environments with varying densities, we randomly removed 25% and 75% of the total number of obstacles. To evaluate the impact of larger obstacles, we removed all obstacles below the median height (0.44 m) ± 1 SD (0.16 m) (Table 3.5). This approach helped evaluate the potential to automate the classification process and assess performance.

3.3.6 SLOPE CLASS CALCULATION

Slope class values are extracted based on the drone-derived DTM. The Sobel operator, which approximates the gradient of the image intensity function, is again used for this calculation. Slope classes are then categorised into the predefined classes (Table 3.6).

While slope classification can include categories based on slope shape (e.g. regular, undulating, terraced, concave, convex) [1], this study did not use this classification. Instead, the slope was calculated from a drone-derived DTM with a 1 m resolution and was classi-

Table 3.4: Ground roughness classification based on obstacle height classes and frequency (H20, H40, H60, H80+ indicate height classes in cm) based on [1].

H20	H40	H60	H80+	Roughness class No. (description)
Infrequent	Isolated	Isolated	Isolated	1 (Smooth)
Moderately frequent	No other classes represented			1 (smooth)
Moderately frequent	Infrequent	Isolated	Isolated	2 (slightly uneven)
Frequent	No other classes represented			2 (Slightly uneven)
Frequent	Moderately frequent	Infrequent	Isolated	3 (uneven)
Moderately frequent	Moderately frequent	Moderately frequent	Infrequent	4 (rough)
Frequent	Frequent	Infrequent	Infrequent	4 (rough)
All surfaces with ground roughness more difficult than that of Class 4				5 (very rough)

Table 3.5: Obstacle count and percentage of the original dataset under different scenarios

Scenario	Obstacle count	% of original
(a) Original dataset	4246	100%
(b) Random removal I	3184	75%
(c) Random removal II	1061	25%
(d) Height-based removal ($\geq$ 0.60 m)	202	5%
(e) Height increased by 16%	4246	100%

Table 3.6: Slope class categories and designations (based on Erasmus [1]).

Class	Slope range (%)	Designation
1	0–11	Level
2	> 11–20	Gentle
3	> 20–30	Moderate
4	> 30–35	Steep 1
5	> 35–40	Steep 2
6	> 40–50	Steep 3
7	> 50	Very steep

fied into slope classes only. This served as complementary data.

3.3.7 Statistical analysis

We used high-resolution drone-derived DTM data to extract statistics on a 3 × 3-meter cell grid. Using a DTM with a 1-meter resolution, we calculated the TRI, VRM, and AR (Table 3.7). For each 3 × 3-meter grid cell, we computed the mean of the 9 underlying 1-meter cells to obtain metrics. This approach allowed us to compare our roughness values with established terrain metrics. We aimed to determine whether these algorithms produce similar results and evaluate whether these metrics are comparable in describing terrain roughness.

Table 3.7: Overview of terrain roughness algorithms and key references

Surface roughness algorithm	Abbreviation	Reference
Terrain ruggedness index	TRI	Riley et al. (1999) [144]
Vector ruggedness measure	VRM	Sappington et al. (2007) [143]
Area ratio	AR	Hobson (1967) [163]

We tested the normality of each variable using the Shapiro–Wilk test. The null hypothesis of normality was rejected at $p < 0.05$. We used the Spearman correlation coefficient (Eq. 3.6) to assess the relationships among the variables.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{3.6}$$

where d_i is the difference between the ranks of corresponding values, and n is the number of observations.

P-values were calculated to determine the statistical significance of the relationships between variables. To ensure comparability of the metrics for plotting, we normalised the TRI, VRM, and AR values using min-max normalisation. This technique scales the data to the [0, 1] range. The normalisation was performed using the following formula (3.7).

$$X' = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{3.7}$$

where X is the original value, X_{min} is the minimum value in the dataset, and X_{max} is the maximum value in the dataset.

3.4 Results

3.4.1 Field and UAV data processing

Data from a preliminary test to extract the most precise data from drone images show that higher-quality settings in point-cloud configurations consistently result in lower measurement errors, with MRE, SD, and RMSE decreasing as quality increases. The test evaluated

all combinations of four quality settings (ultra-high, high, medium, and low) and three depth-filtering modes (aggressive, moderate, and mild) (Table 3.8). The test results, as shown in (Table 3.8), suggested that combining ultra high quality settings with aggressive depth filtering provided the most accurate height measurements, with an MRE of 14.4% (3.3 cm), SD of 3.2 cm, and RMSE of 4.0 cm. RMSE and MRE generally decrease as settings move toward higher, more aggressive configurations, suggesting that these settings are more effective at minimising measurement errors. The RMSE trends across the quality–filtering combinations are summarised in Fig. 3.5. For visual context, Fig. 3.5 includes a 10 cm horizontal reference line to aid interpretation.

Table 3.8: Point cloud quality and depth-filtering combinations with associated error metrics.

Quality	Depth filtering	MRE (%)	SD (cm)	RMSE (cm)
Ultra high	Aggressive	14.4	3.2	4.0
	Moderate	24.3	3.4	6.2
	Mild	28.7	4.2	7.0
High	Aggressive	34.0	7.2	10.7
	Moderate	33.4	5.7	9.3
	Mild	31.8	5.7	9.2
Medium	Aggressive	42.7	7.5	13.0
	Moderate	42.7	7.6	13.1
	Mild	41.9	8.0	13.3
Low	Aggressive	50.0	8.5	15.5
	Moderate	51.0	8.8	15.9
	Mild	49.7	8.6	15.5

Based on these results, the ultra-high-quality settings were selected for subsequent analyses. The resulting point clouds had an approximate density of 8,000 points m^{-2}, supporting more precise estimation of object heights across a range of sizes.

3.4.2 SEGMENTATION PERFORMANCE

The SAM algorithm segmented 86 out of 90 marked obstacles across three study sites, achieving a 95.6% detection rate. The four undetected obstacles, ranging in height from 4 to 15 cm and in circumference from 50 to 110 cm, were likely missed due to complex terrain and dense coverage in the drone imagery.

For the 86 segmented obstacles, the manually measured heights were compared with those derived from the drone's *.las* point cloud data. Fig. 3.6 presents the comparison across three study sites. The regression analysis yields an R^2 of 0.97, indicating that the model explains most of the variability. However, the model underestimates heights, with a mean difference of -2.0 cm (-7.9%) and a standard deviation of 2.6 cm. The RMSE is 2.6 cm

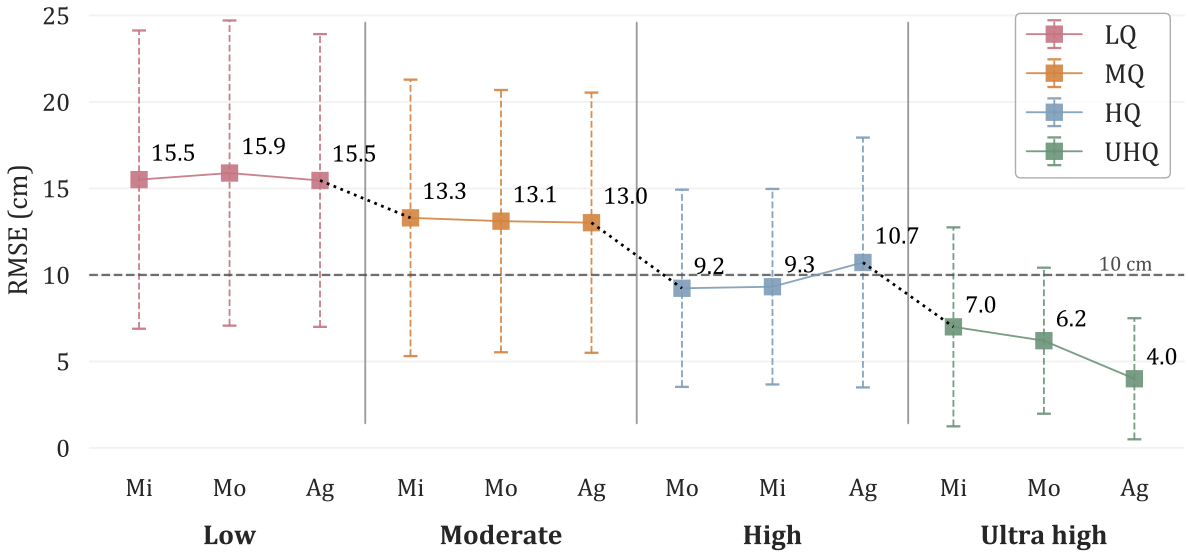


Figure 3.5: Root mean square error (RMSE) across point-cloud quality levels and depth-filtering settings. Points indicate mean RMSE values, error bars represent ±1 standard deviation, and dashed lines connect transitions between quality levels. Vertical separators delineate quality groups (low, moderate, high, ultra-high). The horizontal dashed line indicates a 10 cm reference threshold used for visual comparison. LQ, MQ, HQ, and UHQ denote point-cloud generation settings in Agisoft Metashape corresponding to low, medium, high, and ultra-high quality, respectively.

(10.3%), suggesting modest errors in height predictions. The MRE of 11.4% indicates the average prediction error relative to the actual values. Fig. 3.6(b) shows that predictions are slightly less than the actual values. Predicted heights are closer to actual values for smaller heights, and the underestimation decreases as the actual height increases.

The circumference analysis revealed significant differences between the actual and predicted values. Although the high R^2 of 0.83 indicates a strong linear relationship, the results require careful interpretation. The RMSE value of 73.4 cm (27.3%) and a significant mean difference of 52.4 cm (19.5%) indicate significant prediction errors. The model tends to overestimate circumferences, particularly for larger actual values. While most data points lie within one SD, there are more outliers compared to the height difference plot, indicating higher variability in prediction errors. This issue arises mainly from the algorithm's tendency to merge closely spaced obstacles into a single detected object, leading to increased circumference measurements.

3.4.3 Obstacle-based roughness mapping

Fig. 3.7 shows a distribution of roughness classes across the study area using the original obstacle dataset. The detailed section on the right shows a close-up of a 4x4 cell grid, with each cell measuring 3x3 meters. This grid highlights detected obstacles and their heights in centimetres, with rougher areas containing taller obstacles. The grid-based approach classified the terrain into five roughness classes. The spatial distribution of these classes indicates heterogeneity, with most areas very rough or rough.

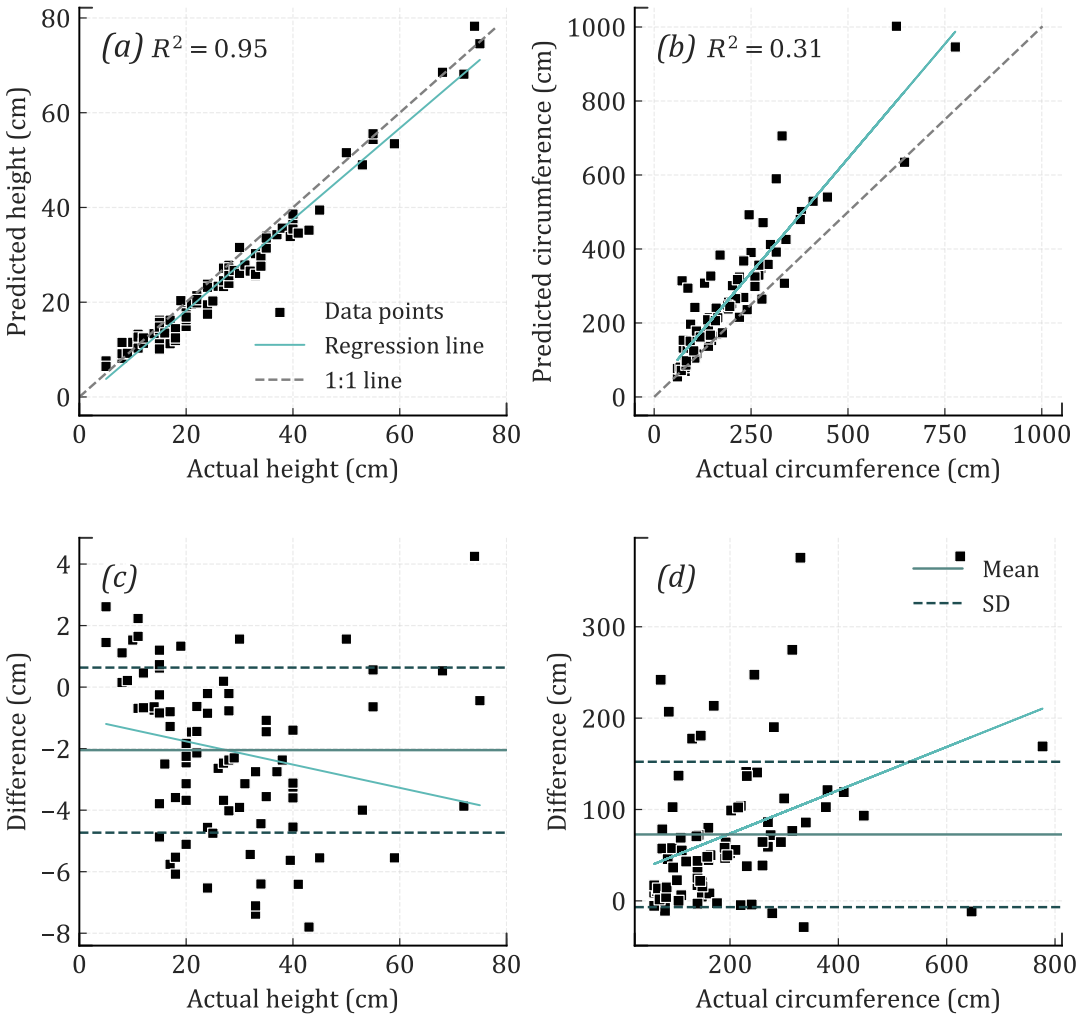


Figure 3.6: Comparison of manually measured and predicted heights and circumferences using drone-derived point cloud data. (a) Linear regression plot for height measurements, where black squares represent data points, the solid blue line indicates the regression fit, and the grey dashed line denotes the 1:1 perfect agreement; (b) difference plot for height measurements showing the mean difference (solid blue line) and ±1 standard deviation (blue dashed lines); (c) linear regression plot for circumference measurements with the same symbology as in (a); (d) difference plot for circumference measurements showing the mean difference (solid blue line) and ±1 standard deviation (blue dashed lines).

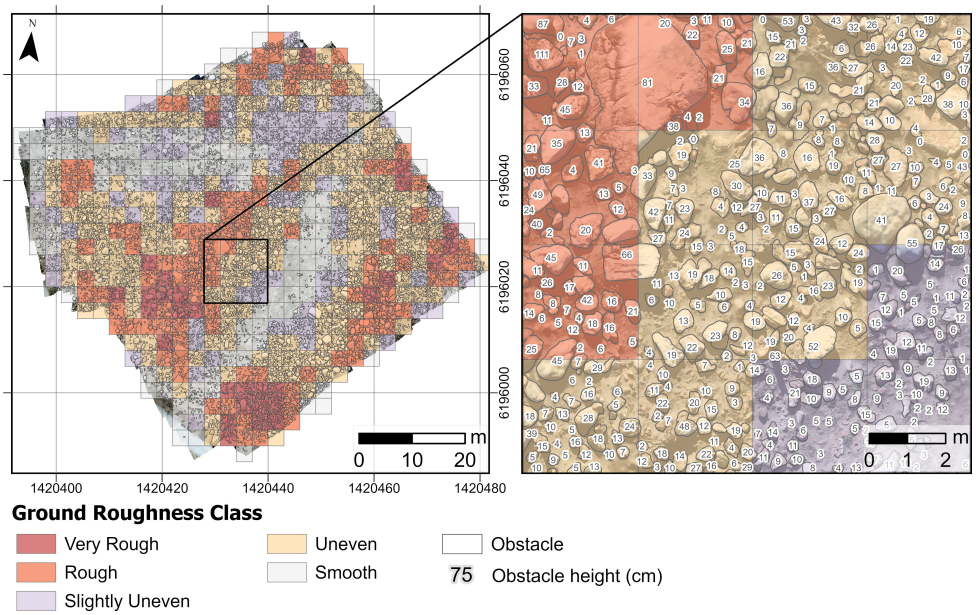


Figure 3.7: Ground roughness classification using the original obstacle dataset. The left map shows the study area's ground-roughness classification using the original set of ob-

Two levels of random obstacle removal were tested to assess the impact on roughness estimation: 25% (Fig. 3.7b) and 75% (Fig. 3.7c). With 25% of the data removed, roughness classification remains stable, retaining recognisable patterns. However, removing 75% of the data leads to significant changes in classification performance, indicating the method's sensitivity to data density. Height-based removal (Fig. 3.7d) eliminates smaller obstacles, leading to a smoother terrain with fewer rough areas. Manipulating obstacle size and distribution affects roughness estimation. Increasing obstacle size generally leads to higher roughness classification, whereas removing or reducing significant obstacles yields smoother terrain.

To address height-detection errors, an additional test increased each obstacle's height by 14.4%. This adjustment resulted in slightly higher roughness classification in some areas (Fig. 3.7d), but the overall roughness distribution remained consistent with the original dataset.

3.4.4 Statistical analysis

The results show that TRI ($\rho = 0.63$, $p < 0.05$) and AR ($\rho = 0.67$, $p < 0.05$) exhibit the strongest correlations with the calculated roughness, indicating a moderate positive relationship. In contrast, VRM ($\rho = 0.24$, $p = 0.13$) shows a weaker correlation (Table 3.9). TRI measures the elevation difference between a centre cell and its surrounding cells, whereas AR is the ratio of surface area to planar area, indicating deviation from a flat plane. The higher correlations between TRI and AR and the calculated roughness likely reflect their sensitivity to elevation differences and surface complexity. In contrast, VRM measures the variability in slope and aspect, making it less sensitive to vertical changes. This could explain why VRM has a weaker correlation with calculated roughness. It might also suggest that the variability obtained by VRM does not align well with the roughness method used in this study.

Table 3.9: Spearman correlation coefficients (ρ) and p-values for mean TRI, VRM, and AR in relation to calculated roughness.

Metric	Spearman correlation ρ	p-value
TRI	**0.63**	**< 0.05**
VRM	0.24	0.13
AR	**0.67**	**< 0.05**

Fig. 3.9 indicates that TRI values increase with higher roughness values. The broader spread of TRI values across higher roughness categories indicates variability in terrain features, suggesting that areas classified as higher roughness (e.g., 4 and 5) exhibit a broader range of TRI values. AR values also increase with higher roughness categories. This trend occurs because AR focuses on the surface's complexity. As the terrain becomes rougher,

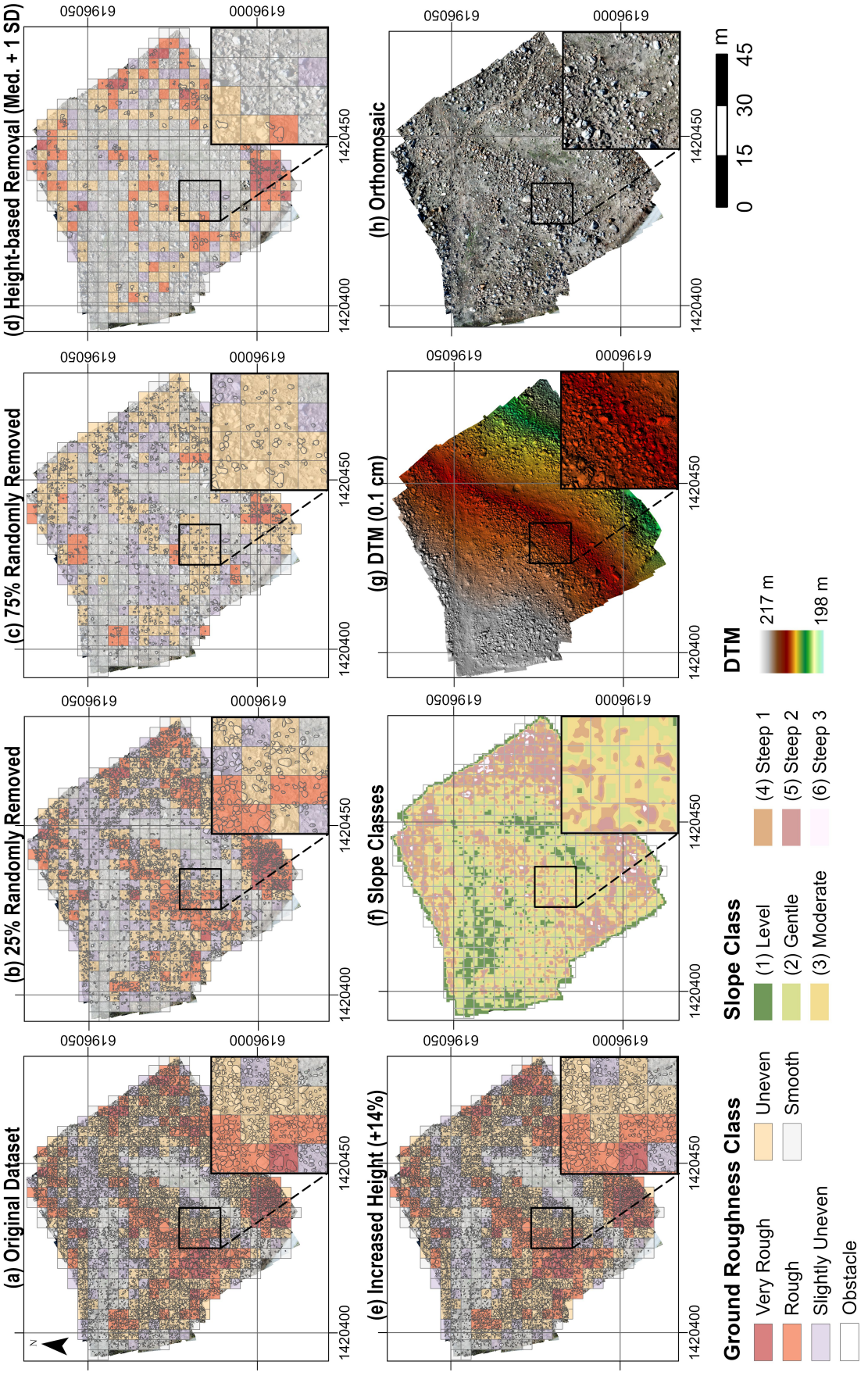


Figure 3.8: Ground roughness and slope classification analysis using a terrain classification system (TCS). (a) Original dataset; (b) 25% randomly removed data; (c) 75% randomly removed data; (d) height-based removal (≥ median + 1 SD); (e) increased height (+16%); (f) slope classes at 1 m; (g) DTM at 0.1 m resolution; (h) original orthomosaic at 0.05 m resolution.

the surface area increases due to greater irregularities, potentially leading to higher AR values. In contrast, VRM exhibits a weaker correlation with roughness and does not show an apparent increase across roughness categories.

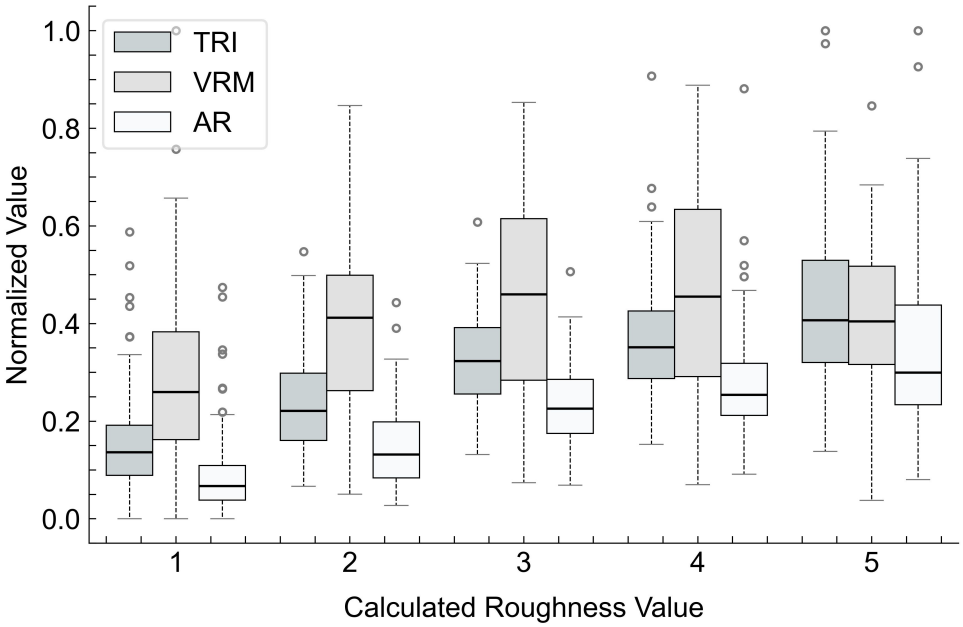


Figure 3.9: Comparison of normalised TRI, VRM, and AR metrics across calculated roughness values. The calculated roughness values are labelled as follows: 1 - smooth, 2 - slightly uneven, 3 - uneven, 4 - rough, and 5 - very rough.

3.5 Discussion

The primary objective of this study was to automate terrain roughness estimation and generate roughness maps for a highly uneven area, using the Jonkershoek timber plantation as the experimental site. While previous research has widely explored methods for surface roughness estimation and characterisation [141, 150, 79], the term "roughness" has been used in various ways, each study offering unique insights. Research has examined soil erosion, landform changes, and terrain features, including road roughness. This study addresses that gap by utilising UAV-derived data and image segmentation to quantify terrain roughness based on the established TCS by [1].

Key findings from this work demonstrate that the SAM algorithm segmented 95.6% of marked obstacles across three study sites, with the model generally providing good predictions of obstacle heights. However, it tends to underestimate heights, with a mean difference of -2.0 cm (-7.9%) and a MRE of 11.4%. This MRE is slightly lower than that obtained from point cloud software comparisons (14.4%) (Table 3.8). While further research is needed to address potential sources of these errors, such as limitations in photogrammetric image processing, insufficient GNSS-located GCPs, or terrain complexity, the model's performance for height prediction remains promising. In contrast, circumference

predictions showed greater variability, with an RMSE of 73.4 cm (27.3%) and a mean difference of 52.4 cm (19.5%), primarily due to the segmentation algorithm merging closely spaced obstacles into a single object. Fine-tuning the segmentation algorithm, such as adjusting proximity settings or employing advanced methods, such as convolutional neural networks (CNNs), could help the model more effectively distinguish nearby objects and reduce errors in circumference predictions.

The original obstacle dataset (Fig. 3.7) provided a detailed roughness map, serving as a baseline for comparison. Reducing dataset density highlighted the method's sensitivity to obstacle density: fewer obstacles yielded smoother terrain classifications. When reducing the dataset by height, the algorithm detected changes and identified areas with sparse yet high-obstacle density.

A test was conducted by increasing the obstacle height by 14.4% above the average estimation error to determine its effect on roughness classification. While the visual comparison suggested higher roughness classifications in some areas, the results did not change significantly, likely because the classification thresholds changed only slightly. Most obstacles remained within the H20 height class, where a 14.4% increase was insufficient to shift many observations into a different class. In addition, the offset moving-window approach may have redistributed and averaged the effects of the height increase across overlapping cells, further reducing its influence on the final roughness classification. The offset moving-window approach recalculates roughness on overlapping grids by applying half-cell offsets in both x and y directions (Fig. 3.4). This ensures that obstacles near cell edges in one grid are also included in neighbouring calculations in the offset grid. Roughness values from the overlapping grids are then averaged, thereby smoothing localised changes in obstacle height and reducing their impact on the final classification.

A comparison between calculated roughness values and established metrics showed that TRI and AR have a stronger relationship with calculated roughness than VRM does. VRM's lower correlation may indicate that its three-dimensional orientation data does not align well with the roughness method used in this study. Conversely, the higher correlations for TRI and AR suggest these metrics better capture aspects of terrain roughness related to elevation differences and surface complexity. However, averaging roughness metrics to 3 × 3 meters likely influenced the results, as these metrics are sensitive to fine-scale terrain variations. A recent study by [46] found that TRI and VRM can yield different results depending on scale and resolution. Despite the terrain's complexity, 3 × 3 meters may be too large for the fine details these metrics aim to capture, leading to averaged metric values. While the correlations between TRI and AR with calculated roughness are moderate, they suggest that these metrics capture key aspects of terrain roughness. Further research is needed to better understand the relationship between these metrics and calculated roughness, potentially by exploring different scales, resolutions, and terrain types. Overall, the trend in TRI and AR values supports our roughness classification, indicating that we detect rougher areas. However, the variability within the values suggests that while our classifi-

cation is generally effective, these metrics may capture finer-scale terrain features more accurately.

The study acknowledges several limitations. While orthomosaic offers a clear view of the terrain in our case, detecting obstacles under dense canopies remains challenging. UAV imagery is advantageous due to its availability, ease of data collection, and cost-effectiveness, making it useful for open landscapes and sparsely vegetated areas. However, in forest environments where canopy cover reduces visibility, technologies such as TLS, ALS, and mobile LiDAR can provide significantly higher resolution and greater under-canopy data coverage [164]. Although these technologies are becoming more accessible, their high costs and specialised equipment requirements still pose barriers to widespread use. This study establishes a UAV-based roughness estimation approach as a cost-effective alternative to LiDAR, providing a baseline methodology for future refinement. Drones offer a straightforward, cost-effective, and accessible means to assess terrain roughness, serving as a practical starting point before integrating more complex and expensive technologies such as LiDAR. However, to improve roughness estimation in dense environments, future research should explore point cloud classification methods that leverage DEMs and high-quality point clouds from TLS or ALS. Given LiDAR's ability to penetrate canopy and provide high-resolution ground-level detail, integrating it into roughness estimation would improve its applicability across a broader range of environments. This would also enable more accurate obstacle detection and extraction.

Advanced deep-learning algorithms trained specifically for terrain and obstacle detection, though requiring substantial training data, could also be applied to high-resolution point clouds. Previous studies have explored the use of PointNet++ and other deep learning techniques to classify forest attributes and to provide accurate tree-structure information. For instance, PointCNN achieved up to 87.0% accuracy using LiDAR points, laser intensity, and multispectral features for tree detection and canopy cover estimation [165, 166]. Combining UAV and TLS data for estimating forest tree metrics was also evaluated by [138], achieving 97.8% accuracy for diameter at breast height and total tree height estimates for broadleaf trees. A similar approach could also be applied to terrain features. Including slope, soil moisture, and roughness measurements in a TCS could provide a better understanding of surface conditions. While slope and its shape can be extracted from point clouds, soil moisture estimation often requires satellite or other remote sensing data. Using Sentinel-1 and Sentinel-2 data with deep learning models has been effective for estimating soil moisture, achieving high correlation (0.80) and low RMSE (0.04 m) [167].

A further limitation of this study relates to the representativeness and diversity of the training data used to adapt the segmentation workflow. Although the obstacle detection approach was based on a pre-trained SAM model, it was customised using a relatively limited number of manually labelled samples collected under specific site conditions. Consequently, segmentation performance may be constrained when applied to environments that differ substantially from the experimental setting and may also be influenced by image-

acquisition characteristics, such as illumination, shadowing, sensor angle, and flight timing. These sources of variability can affect feature extraction and classification performance. Expanding the training dataset to include a broader range of terrain types, vegetation structures, and acquisition conditions would improve model robustness and transferability across environments that require such adaptation.

The next step is to combine point cloud data with deep learning to improve terrain-roughness assessment. Expanding the study from the experimental setup to forest environments will help generalise the method. A key area for future development is the ability to perform roughness estimation in real or near real time to support operational decisions, such as pre-harvest planning, route optimisation, and transportation. Current efforts focus on using TLS as a proxy for machine-mounted sensors and on adapting the existing roughness estimation methodology to use LiDAR data rather than UAV imagery. Additionally, using accurate GCP placement and positioning will be crucial for ensuring data reliability and geospatial alignment. This would improve accuracy, adaptability, and real-time terrain analysis.

However, scaling this method to larger and more complex areas introduces additional challenges. Processing high-resolution point clouds over large spatial extents can be computationally demanding. It may benefit from high-performance computing resources (e.g., increased RAM and, depending on the workflow, GPU acceleration) to support efficient point-cloud generation and analysis. Workflow optimisations such as tiling, batching, and parallel processing are therefore essential to maintain practical runtimes. In addition to computational efficiency, transparent and automated classification workflows are needed to streamline data processing and improve scalability across diverse datasets and acquisition methods.

In addition to computational efficiency, transparent and automated classification workflows are required to streamline data processing and ensure scalability across diverse datasets, regardless of the data-collection method.

3.6 Conclusion

This study has shown that UAV imagery can be used to automatically assess and categorise terrain surface roughness, creating detailed roughness maps. The algorithm adapted well to changes in obstacle height and incidence, highlighting the potential of using remote or proximally sensed point clouds, regardless of their origin, to describe and characterise terrain. Furthermore, our findings suggest that, although TRI and AR metrics exhibit similar trends with our calculated roughness values, they may not capture roughness in the same way. These metrics reflect key aspects of elevation and surface complexity, but their variation with elevation may differ from the patterns observed with our method.

Future research could focus on using point cloud data to provide more robust three-dimensional insights and improve terrain feature detection. Fully automating a TCS that in-

cludes roughness, slope, and soil moisture can enhance decision-making in forestry and environmental management. This automation would not only improve terrain assessments but also help set realistic expectations for extraction performance and costs, thereby improving operational efficiency. Additionally, incorporating more datasets and applying these classifications to trafficability and forestry path planning could further streamline operations and reduce costs.

3.7 DATA AVAILABILITY

The datasets generated and analysed in this study are available in the University of Padova Research Data repository at https://researchdata.cab.unipd.it/1723/.

The corresponding published paper is available at https://www.tandfonline.com/doi/full/10.1080/14942119.2025.2480009.

Chapter 4

Detecting ground-level obstacles using high-density point cloud

4.1 Abstract

Reliable identification of near-ground obstacles is critical for terrain characterisation and trafficability assessment in forest operations, particularly in environments where canopy cover limits aerial observation. This chapter investigates the feasibility of detecting ground-level obstacles using high-density terrestrial laser scanning (TLS) point clouds, with emphasis on point-level semantic separability rather than object-level quantification. Using a post-fire plantation environment as a test case, TLS data were manually annotated into three operationally relevant classes: ground, low vegetation, and rock, and used to train a supervised semantic segmentation model based on the PointNet++ architecture. Model performance was evaluated on spatially independent test blocks using standard point-wise accuracy and intersection-over-union metrics. The effect of including TLS intensity as an additional input feature was also assessed. Results show that high-density TLS data enable consistent semantic separation of near-ground classes, with global accuracies of approximately 0.87–0.90 and mean IoU values of 0.76–0.78 on held-out test data. Incorporating intensity information yielded modest but systematic improvements across all classes, particularly for discriminating rock surfaces from surrounding terrain. Low vegetation remained the most challenging class, reflecting structural heterogeneity and diffuse boundaries with adjacent surfaces. The findings demonstrate that TLS provides sufficient geometric detail to support supervised segmentation of ground-level obstacles under sparse-canopy conditions. However, the analysis is limited to a single high-density TLS dataset and focuses on segmentation feasibility rather than full-scene inference or obstacle-level validation. The chapter therefore serves as an experimental and exploratory step toward integrating TLS-based obstacle mapping with the UAV-based roughness assessment presented in Chapter 3, highlighting both the potential and current limitations of point-cloud-based approaches for near-ground terrain characterisation.

4.2 INTRODUCTION

Forests are complex and dynamic environments in which terrain variability poses challenges for both operational planning and machine-based tasks. Terrain features, such as slope variation, surface roughness, and dense vegetation, can significantly affect machine stability, limit accessibility, and increase operational risk. In fact, terrain-related limitations are among the primary factors contributing to delays, increased fuel consumption, and equipment wear in forestry operations. Studies have shown that factors such as steep slopes [86, 146], saturated soils [168], and rough trail surfaces [146, 169] can significantly reduce machine productivity, increase energy demands, and accelerate wear. In some cases, fuel use has been observed to increase by more than 100% under extreme terrain conditions [146], highlighting terrain as one of the most influential variables in forestry operations. Quantifying these characteristics is therefore essential for effective operational planning, risk mitigation, and trafficability assessment.

Historically, terrain accessibility has been evaluated using terrain classification systems (TCS), which categorise land based on key physical features [1, 41, 43, 40, 44]. Typical TCS frameworks combine a) ground conditions related to soil bearing capacity and moisture, (b) ground roughness associated with discrete obstacles such as rocks or depressions, and (c) slope and topographic form [1, 41]. While effective for operational decision-making, these assessments are often performed manually and rely on expert judgement, making them time-consuming, subjective, and difficult to standardise across spatial scales [1].

To reduce subjectivity and improve spatial detail, recent research has explored automated terrain evaluation using remote sensing products such as UAV photogrammetry and surface models. In Chapter 3, UAV-derived data were used to approximate roughness-related components of TCS with high spatial resolution. However, UAV and image-based approaches remain constrained in densely forested environments where canopy occlusion limits ground visibility and reduces the reliability of near-ground obstacle detection.

From an operational perspective, discrete ground-level obstacles such as rocks, boulders, and dense low vegetation can have a disproportionate impact on machine mobility, stability, and energy demand. Obstacles obscured beneath shrubs or regenerating vegetation can induce vibration, slippage, or route deviations and may increase fuel consumption and operational costs. Reliable mapping of near-ground obstacles therefore remains a key gap in terrain characterisation in forest environments with limited visibility.

LiDAR-based sensing provides a robust alternative for capturing 3D terrain structure under vegetation. In particular, terrestrial laser scanning (TLS) offers very high point density and near-surface geometric detail, enabling improved representation of rocks, stumps, shallow depressions, and microtopography compared with aerial or image-based methods [61, 170]. TLS has been widely applied in forestry to characterise both above- and below-canopy structure, including understorey complexity and microtopographic surfaces that

are difficult to observe with nadir-view remote sensing [171]. TLS-derived surface products have also been used to quantify terrain deformation and fine-scale changes in roughness caused by forest operations. For example, millimetre-level surface change and volumetric metrics have been reported for skid-trail disturbance and degradation [169].

For mobility and routing applications, TLS has also been used to derive near-ground obstruction metrics [62], quantifying near-ground obstruction density and height using voxel-based occupancy derived from TLS, demonstrating that low vegetation can substantially influence traversability predictions and that TLS can serve as a high-detail reference for evaluating coarser LiDAR-based terrain products.

At the same time, advances in point-cloud-based machine learning have enabled semantic segmentation directly on irregular 3D point clouds. Deep learning architectures such as PointNet++ [172] and RandLA-Net [173] can assign point-wise labels while preserving fine-scale geometric structure, and these approaches are increasingly adopted in forestry LiDAR research. Recent work has focused on general forest-scene understanding across sensors and forest types, including semantic separation of foliage, wood, ground, and lower vegetation using dedicated deep learning frameworks [174]. In parallel, the availability of large annotated benchmark datasets has accelerated method development and comparability, for example, the ground-based SegmentedForests dataset (>920 million labelled points across multiple forest plots and sensors) [175] and the FOR-instance UAV laser scanning benchmark for semantic and instance segmentation [176]. These developments demonstrate that deep learning segmentation of forest point clouds is now feasible at scale. However, most efforts focus on forest structural components (e.g., stems, crowns, foliage/-wood separation) rather than on discrete ground-level obstacles [177]. The applicability of these methods to near-ground obstacle detection remains underexplored, particularly in environments where low vegetation obscures rock surfaces and class boundaries are diffuse.

This chapter explores the feasibility of detecting ground-level obstacles from high-density TLS point clouds using supervised semantic segmentation. Building on the limitations identified in Chapter 3, where UAV-based approaches were constrained by canopy obstruction and limited near-ground visibility, the study evaluates whether TLS provides sufficient detail to separate three operationally relevant classes: rock, low vegetation, and ground in a post-fire forest environment.

The analysis is preliminary and focuses on class-level separability and the practical performance of a baseline segmentation pipeline, including the contribution of TLS intensity as an additional input feature. Rather than detailed obstacle reconstruction or direct quantification of trafficability, the emphasis is on segmentation performance and interpretability as a first step toward applications such as machine route planning and terrain assessment, where near-ground obstacles can affect delays, fuel consumption, and operating costs.

4.3 MATERIALS AND METHODS

This section describes the materials and methods used to evaluate ground-level obstacle separability from high-density TLS point clouds, including TLS acquisition, preprocessing and georeferencing, manual annotation, data augmentation, supervised semantic segmentation with PointNet++, and evaluation on spatially independent test blocks (Fig. 4.1).

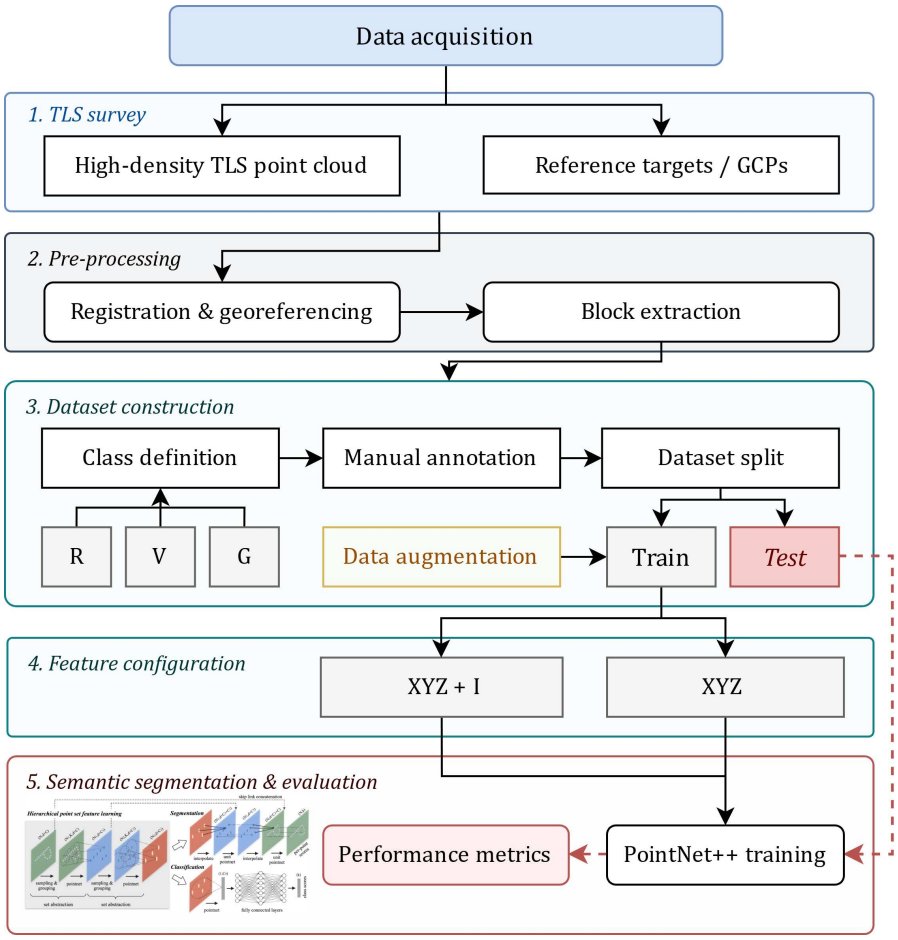


Figure 4.1: Experimental workflow for ground-level obstacle detection using high-density TLS data. TLS point clouds and reference targets are used for scan registration and georeferencing, followed by block extraction. Point-cloud blocks are manually annotated into terrain classes and split into training and test subsets, with data augmentation applied exclusively to the training data. Semantic segmentation is evaluated under two alternative input feature configurations (XYZ and XYZ + intensity) using an identical PointNet++ architecture, and performance is assessed on the held-out test set.

4.3.1 STUDY AREA

The study was conducted within the Jonkershoek timber plantation in the Western Cape province of South Africa (33'19.61" S, 18'42.24" E). The study site is the same post-fire plantation landscape described in Chapter 3. The region has a Mediterranean climate with hot, dry summers and mild, wet winters [156], and a high incidence of wildfires, particularly between December and May [157, 158].

A few years before this study, a forest fire substantially altered the site's vegetation structure. During data collection, vegetation remained sparse, dominated by shrubs approximately 1.5 to 2 m tall, with exposed rocks of varying sizes across the terrain. Under typical pre-disturbance conditions, the stand would have been harvested using motor-manual felling, with tree-length extraction by cable skidder. Field conditions are illustrated in Fig. 4.2.

Figure 4.2: Overview of the study area located in a post-fire pine plantation landscape in the Jonkershoek Nature Reserve, Western Cape, South Africa. (a–b) Typical terrain and vegetation conditions during data collection, showing exposed rocks and dense low-vegetation regeneration. (c) The RIEGL VZ-2000i terrestrial laser scanner used in this study.

4.3.2 TLS data collection

TLS data were acquired using a RIEGL VZ-2000i, a time-of-flight terrestrial laser scanner designed for long-range, high-density 3D mapping. The instrument provides a 360° horizontal and 100° vertical field of view and records dense point returns suitable for near-ground terrain characterisation (Table 4.1). The TLS instrument used is shown in Fig. 4.2c.

The data were collected during two field campaigns in April and May 2025. The April 2025 campaign covered the area corresponding to Plots 1 and 2 (Chapter 3; Fig. 3.2), while the May 2025 campaign covered the area corresponding to Plot 3 (Chapter 3; Fig. 3.2).

Scans were acquired from tripod-mounted positions spaced approximately 15 m apart, with a consistent scanner height of 2 m. Multiple overlapping scan positions were used to increase point density and reduce occlusions caused by rocks and low vegetation. In total, 27 scan positions were acquired during the April 2025 campaign and 30 scan positions during the May 2025 campaign.

A total of 13 ground control points (GCPs) were surveyed using a high-precision GNSS receiver and a Trimble TSC3 controller. These GCPs were used as control targets for subsequent georeferencing of the registered TLS point clouds.

Table 4.1: Technical specifications of the RIEGL VZ-2000i terrestrial laser scanner.

Parameter	Value
Field of view (FOV)	360° (H) × 100° (V)
Resolution	7201 × 2082 scanlines
Maximum range	2000 m
Scan time	29 s per scan
Beam divergence	0.3 mrad
Measurement rate	1.2 MHz (up to 14,992,482 points per scan)
Wavelength	Infrared
Accuracy	5 mm
Repeatability	3 mm

4.3.3 TLS preprocessing and georeferencing

All TLS scans were processed using RIEGL RiSCAN PRO v2.20.1 [178]. Each scan position was processed individually and exported in LAS format. The LAS files were then imported into CloudCompare for preprocessing and georeferencing. In CloudCompare, scan positions from each campaign were merged into a single point cloud scene using the *Merge multiple clouds* tool.

The merged point cloud was georeferenced using the surveyed GCPs. An initial coarse alignment was performed using the CloudCompare *Align two point clouds* tool based on manually selected corresponding GCP locations. The alignment was subsequently refined using iterative closest point (ICP) registration. The final point clouds were referenced to the WGS 84 / UTM Zone 32S coordinate system. To reduce isolated outliers, the CloudCompare noise filter was applied with a radius of 0.5 m, removing points with insufficient local neighbourhood support.

4.3.4 Manual annotation and class definition

Following preprocessing and georeferencing, the TLS point cloud was spatially subset into six non-overlapping blocks of 35 m × 35 m distributed across the scanned area. Blocks were selected to capture representative variability in terrain and vegetation conditions and to maximise scene diversity for subsequent analysis.

Manual annotation was performed in CloudCompare using the *Segment* tool. Within each block, points belonging to individual features were interactively segmented and assigned to one of three classes: *rock*, *low vegetation*, and *ground*. Segmented objects belonging to the same class were merged into a single class-specific point cloud (i.e., one cloud containing only rock points, one containing only vegetation points, and one containing only ground points).

A categorical label was then stored in the CloudCompare *Classification* scalar field using *Add Classification SF*, with class codes assigned as rock = 1, low vegetation = 2, and

ground = 3. Finally, the class-specific point clouds were merged back into a single labelled point cloud using the *Merge two clouds* tool, preserving the *Classification* field for subsequent model development and evaluation. All annotations were performed by the author to ensure consistent interpretation of class boundaries.

4.3.5 Data augmentation strategy

To increase the amount of labelled training data and improve robustness to sampling variability, data augmentation was applied to the manually annotated TLS blocks. All augmentations preserved the original class labels. Transformations did not alter the class identity of points (ground, low vegetation, rock). However, it introduced additional geometric variability expected in TLS acquisitions, such as differences in scene orientation, local measurement noise, and partial occlusion.

Four augmentation operations were applied (Table 4.2): random rotation around the vertical axis, additive jittering noise, axis flipping, and random point dropout. The resulting augmented point clouds were used as additional training samples in the supervised semantic segmentation experiments described in the following subsection. Augmentation was implemented in Python, and the augmented point clouds were exported to LAS format for subsequent model training.

Table 4.2: Point cloud data augmentation strategies with their purposes, key parameters, and implementation notes.

No.	Augmentation	Purpose	Key parameters	Number of variants
1	**Z-axis rotation**	Introduces orientation variability by rotating the point cloud around Z.	`angle`: Random in [0, 360] degrees	6 variants per scene
2	**Jittering**	Adds noise to simulate measurement inaccuracies.	`std`=0.01, `clip`=0.05 (Gaussian noise, clipped)	6 jittered versions per scene
3	**Axis flipping**	Mirrors scenes along X and/or Y axes.	`flip_x`, `flip_y`: True/False	3 versions: flipX, flipY, flipXY
4	**Point dropout**	Simulates occlusion or incomplete data.	`dropout_rates`={0.1, 0.2, 0.3, 0.5}	Applied globally (not per class)

4.3.6 Supervised semantic segmentation framework

Supervised semantic segmentation was performed to classify TLS points into three classes: ground, low vegetation, and rock. Model development and experiments were conducted in

MATLAB R2025b (MathWorks) [179]. All model training and evaluation were performed on a workstation with the technical specifications listed in Table 4.3.

Table 4.3: Technical specifications of the workstation used for experiments.

Parameter	Value
CPU	Intel Xeon W-2295, 18 cores / 36 threads
GPU	NVIDIA RTX 4000 Ada Generation
Dedicated GPU Memory	20 GB
Total Available GPU Memory	83.6 GB (including 63.6 GB shared)
System RAM	128 GB DDR4

The PointNet++ architecture [172] was used as the baseline model for its ability to learn hierarchical features directly from unordered point clouds. The network was trained using manually annotated TLS blocks (Section 4.3.4) and augmented samples (Table 4.2).

Each point was represented by its 3D coordinates (X, Y, Z). Two input configurations were evaluated: (a) geometric features only, using (X, Y, Z), and (b) geometric features plus TLS intensity, using (X, Y, Z, I), where I denotes the recorded return intensity. This comparison was included to assess whether intensity provides complementary information for separating near-ground classes beyond that provided by geometry alone, while holding all other settings constant.

To provide a fixed-size input to PointNet++, each training and test sample was constructed by sampling a fixed number of points per block (e.g., $N_p = 8192$ points). Unclassified points (label 0) were excluded from training and evaluation.

The six labelled 35 m × 35 m blocks were split at the block level to avoid spatial leakage: four blocks were used for model training and two blocks were held out for evaluation. Data augmentation was applied only to the training blocks to generate additional training samples. No augmentation was applied to the held-out evaluation blocks.

Model training was run for up to 200 epochs. An early stopping criterion was applied to prevent overfitting, terminating training when validation performance did not improve for a predefined number of consecutive epochs. The complete set of training hyperparameters is reported in Table 4.4.

4.3.7 Evaluation metrics

Model performance was evaluated on the held-out test blocks using standard point-wise classification metrics. Overall performance was assessed using global accuracy (GA; Eq. 4.1), mean accuracy (MA; Eq. 4.2), mean intersection over union (M-IoU; Eq. 4.3), and weighted IoU (W-IoU; Eq. 4.4). Per-class performance was evaluated using per-class accuracy (A_c; Eq. 4.5) and per-class IoU (IoU_c; Eq. 4.6) for the ground, low vegetation, and rock classes.

Table 4.4: Training hyperparameters and optimisation settings used for PointNet++ model training in MATLAB R2025b (MathWorks) [179]

Setting	Value
Optimizer	Adam
Initial learning rate	5×10^{-4}
L2 regularization	0.01
Max epochs	200
Mini-batch size	32
Shuffle	every epoch
Learning-rate schedule	piecewise
Learning-rate drop factor	0.5
Learning-rate drop period	20 epochs
Gradient decay factor (β_1)	0.9
Squared-gradient decay factor (β_2)	0.999
Execution environment	auto (CPU/GPU)
Checkpointing	enabled (saved during training)
Training monitoring	training-progress plot

$$GA = \frac{\sum_c TP_c}{N}. \tag{4.1}$$

$$MA = \frac{1}{C}\sum_c A_c. \tag{4.2}$$

$$M\text{-}IoU = \frac{1}{C}\sum_c \mathrm{IoU}_c. \tag{4.3}$$

$$W\text{-}IoU = \sum_c \left(\frac{n_c}{N}\right)\mathrm{IoU}_c. \tag{4.4}$$

$$A_c = \frac{TP_c}{TP_c + FN_c}. \tag{4.5}$$

$$\mathrm{IoU}_c = \frac{TP_c}{TP_c + FP_c + FN_c}. \tag{4.6}$$

Here, c indexes the classes and C is the total number of classes. TP_c, FP_c, and FN_c denote the numbers of true positives, false positives, and false negatives for class c, computed from point-wise comparisons between predicted and reference labels. N is the total number of evaluated points, and n_c is the number of reference points belonging to class c.

4.4 Results

4.4.1 Labelled TLS dataset

The manually annotated dataset consisted of six non-overlapping TLS blocks of 35 m × 35 m distributed across the study area. The labelled point cloud comprised more than 16 million points, classified into three classes: rock (Class 1), low vegetation (Class 2), and ground (Class 3).

Across the six labelled blocks, the mean point density was approximately 11 700 points m^{-2}. Class-specific mean densities were highest for rock surfaces (approximately 6 100 points m^{-2}), followed by low vegetation (approximately 4 100 points m^{-2}), and ground (approximately 1 500 points m^{-2}).

Fig. 4.3 illustrates the relative class composition of each labelled block. Rock points accounted for the majority of the total point count in most blocks, whereas the proportions of low vegetation and ground varied substantially across plots, reflecting spatial heterogeneity in terrain structure and obstacle distribution.

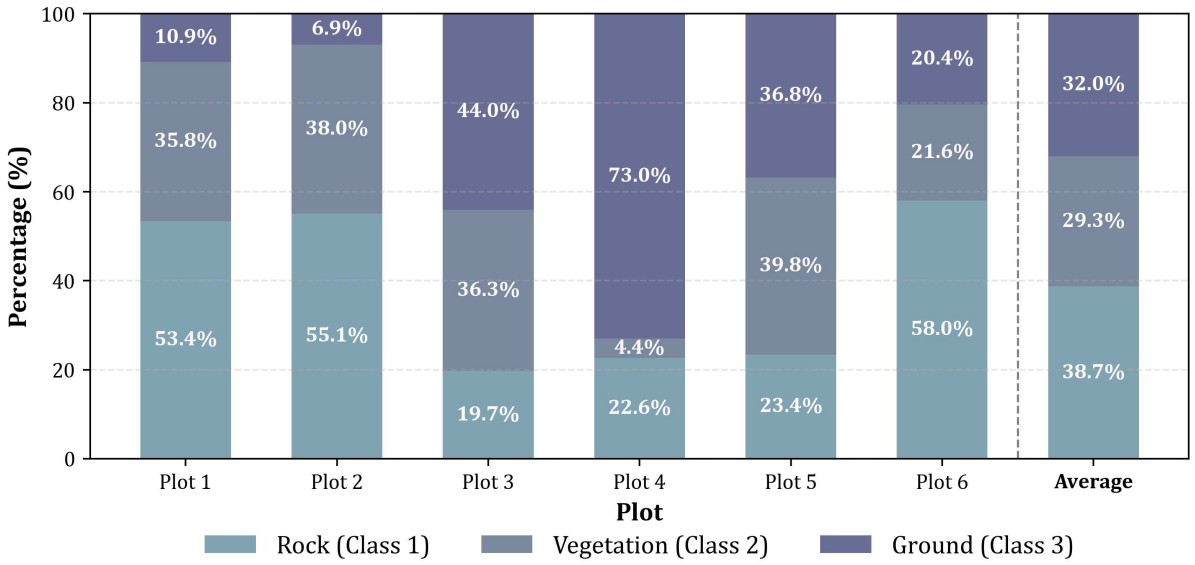


Figure 4.3: Class composition of the manually annotated TLS blocks. Bars show the percentage of points assigned to ground, low vegetation, and rock for each labelled block; the final bar indicates the mean class distribution across all blocks.

In addition to geometric density and class composition, the labelled dataset also exhibited class-dependent TLS intensity behaviour. Fig. 4.4 shows intensity distributions for manually annotated points by class. Rock and ground have higher central intensity values than low vegetation, although the distributions overlap substantially.

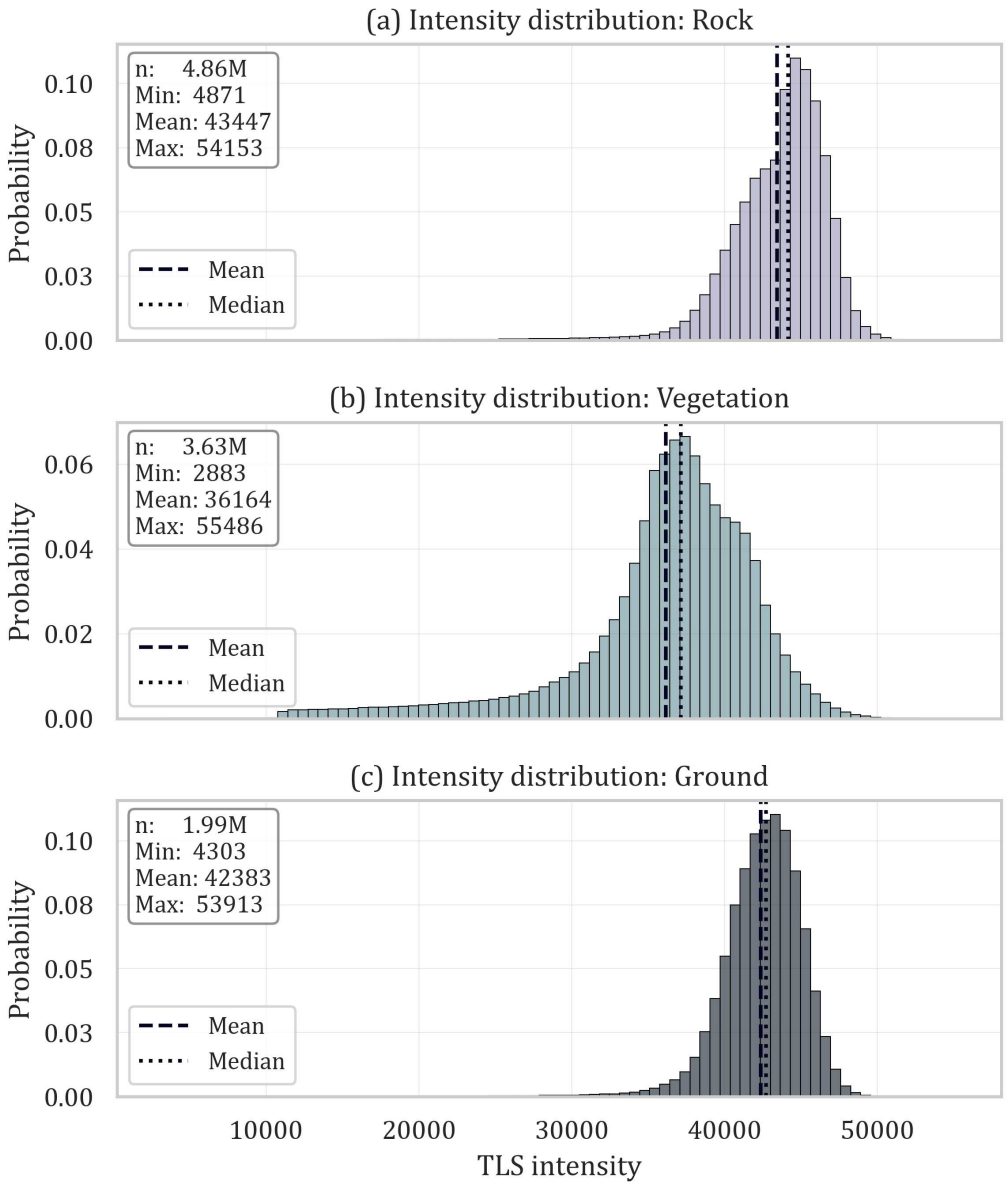


Figure 4.4: TLS intensity distributions for manually annotated points belonging to (a) rock, (b) low vegetation, and (c) ground classes. Distributions are derived only from labelled TLS blocks. Dashed and dotted vertical lines indicate the mean and median intensity values, respectively.

4.4.2 EFFECT OF TRAINING DATA AUGMENTATION

Data augmentation substantially increased the amount of labelled training data by generating multiple transformed variants of each manually annotated TLS block (rotation, jittering, flipping, and point dropout), while preserving the original class labels. As a result, the total number of labelled points used for training increased from 16.3 million to 478.4 million ($\approx 29\times$). Mean point density increased from 11,724 to 12,893 points m^{-2} ($\approx +10\%$), indicating a modest densification of the training data after augmentation and preprocessing.

Class-specific point densities changed after augmentation, with rock remaining nearly unchanged ($\approx$9.3%), low vegetation increasing moderately ($\approx$+8.8%), and ground increas-

Table 4.5: Overall segmentation performance for XYZ vs. XYZ+intensity (higher is better). Absolute improvement is reported as Δ (XYZ+I − XYZ), with relative improvement in parentheses.

Input	Global acc.	Mean acc.	Mean IoU	Weighted IoU
XYZ	0.87	0.87	0.76	0.77
XYZ+intensity	0.90	0.90	0.78	0.78
Improvement (Δ)	+0.03 (+3.33%)	+0.03 (+3.25%)	+0.03 (+3.50%)	+0.01 (+1.17%)

ing substantially (≈+35.9%). This indicates that augmentation increased the effective training data volume and shifted the class density distribution toward the ground, in addition to introducing geometric variability relevant to TLS acquisition conditions.

4.4.3 Semantic segmentation performance

Table 4.5 summarises test-set semantic segmentation performance for the two input feature configurations (XYZ vs. XYZ+intensity). Using only XYZ coordinates, the PointNet++ model achieved a global accuracy of 0.87 and a mean IoU of 0.76. When TLS intensity was included as an additional input feature, performance improved consistently across all aggregate metrics, with global accuracy increasing to 0.90 and mean IoU to 0.78.

Per-class performance metrics are reported in Table 4.6. Ground points were classified with the highest accuracy, reflecting the relatively planar and continuous nature of the ground surface. Rock classification benefited most from the inclusion of intensity information, with IoU increasing from 0.79 to 0.82. Low vegetation exhibited the lowest accuracy and IoU across both feature configurations, consistent with its structural variability and diffuse boundaries with adjacent classes. Fig. 4.5 shows representative point-wise predictions produced by the PointNet++ model, illustrating spatially coherent class regions and the separation between ground surfaces, rocks, and low vegetation.

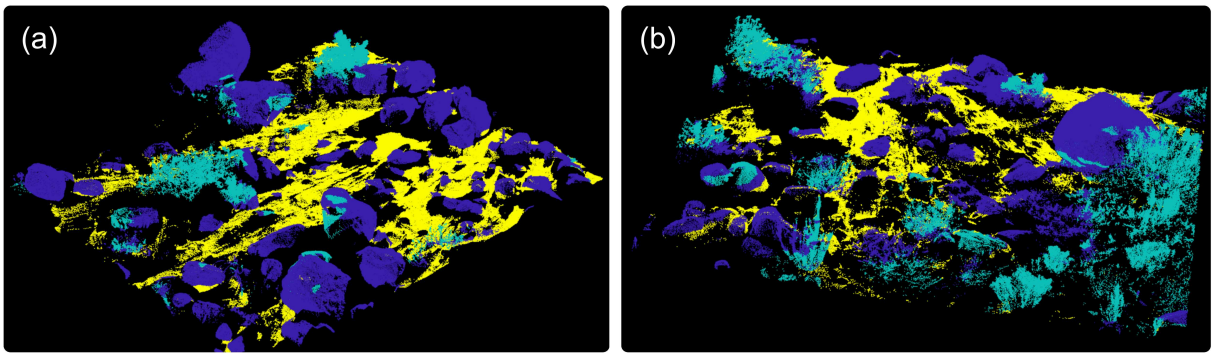


Figure 4.5: Example semantic segmentation results produced by the PointNet++ model using XYZ coordinates. Ground points are shown in yellow, low vegetation in cyan, and rock in purple. Panels (a) and (b) show two representative viewpoints illustrating the spatial coherence of predicted classes and the separation between ground surfaces, rocks, and low vegetation.

Table 4.6: Per-class IoU for XYZ vs. XYZ+intensity (higher is better). Improvements are Δ (XYZ+I − XYZ) with relative improvement in parentheses.

	Rock	Vegetation	Ground
IoU (XYZ)	0.79	0.76	0.72
IoU (XYZ+I)	0.82	0.79	0.78
Improvement (Δ)	+0.03 (+3.77%)	+0.03 (+3.58%)	+0.06 (+8.07%)

4.5 Discussion

This chapter evaluated the feasibility of separating ground, low vegetation, and rock classes from high-density TLS point clouds using a baseline supervised semantic segmentation workflow. Results obtained on held-out test blocks indicate that point-level semantic separation is achievable under the investigated post-fire conditions, with global accuracy of approximately 0.87–0.90 and mean IoU of approximately 0.76–0.78. Visual inspection of predicted segmentation maps further indicates spatially coherent class regions, suggesting that TLS-based semantic segmentation can capture meaningful near-ground terrain structure at local scales.

Including TLS intensity as an additional input feature led to consistent, though moderate, performance improvements across all evaluated metrics. Class-wise intensity distributions derived from the labelled data show shifts in central tendency between rock and ground relative to low vegetation, which provides a plausible explanation for the observed performance gains when intensity information is available. However, the substantial overlap between class distributions indicates that intensity (Fig. 4.4) alone does not provide a robust decision rule and should be interpreted as complementary to geometric information rather than as a standalone discriminator.

A further consideration is the transferability of intensity as an input feature. TLS intensity is typically instrument- and acquisition-dependent and can vary with range, incidence angle, and scanner settings. Without explicit radiometric normalisation or calibration, intensity-driven performance gains observed in this dataset may not generalise across campaigns or sensors. Future work should therefore document the intensity preprocessing and evaluate whether the improvements persist across different acquisition geometries and platforms.

Although intensity improved segmentation performance, further gains are likely to be achieved by enriching the input representation with explicit local geometric descriptors. While PointNet++ can learn geometric patterns directly from raw coordinates, providing neighbourhood-based shape features may improve separability in near-ground scenes characterised by diffuse boundaries, occlusions, and mixed returns. Potential descriptors include verticality, planarity, curvature or normal change rate, and eigenvalue-based features derived from the local covariance structure of neighbouring points.

This is consistent with prior LiDAR literature, which shows that neighbourhood-based

eigenvalue and normal-derived descriptors can be strongly discriminative when raw XYZ data alone is insufficient. For example, [180] developed an automated TLS leaf–wood separation workflow that explicitly computes multiple eigenvalue-based geometric features from local neighbourhoods and uses them for point-wise classification. In a related TLS tree-structure pipeline, [181] similarly relies on multi-scale neighbourhood features derived from local covariance eigenvalues, including linearity and planarity, and a normal-based verticality measure, noting that the apparent geometric character of points (linear, planar, volumetric) can change with neighbourhood scale.

Eigenvalue-based features have also proved helpful for point cloud classification beyond individual trees. [182] tested a set of covariance-derived descriptors (e.g., omnivariance, eigenentropy, planarity, sphericity, surface variation, anisotropy, linearity, and verticality) and used class-wise histograms to identify features that best separate vegetation from terrain, finding evident differences for measures such as eigenentropy and omnivariance. These results align with expected near-ground class geometry, where ground points are usually locally planar with stable normals, low vegetation shows more variable normals and vertical structure, and rock surfaces often have stronger curvature and anisotropy around edges and boulder faces. Adding a small set of such neighbourhood-based descriptors alongside XYZ and intensity may therefore help reduce confusion between low vegetation and rock in partially occluded areas.

However, the experimental setup also has several significant limitations. All results are derived from a single high-density TLS dataset (mean approximately 11 ,700 points m^{-2}) acquired in a post-fire plantation environment. Model performance under lower point densities (e.g., 1,000–3,000 points m^{-2}), different acquisition geometries, or alternative platforms such as mobile or airborne laser scanning (MLS/ALS) remains untested. Consequently, the reported metrics should be interpreted as a site- and sensor-specific feasibility demonstration under favourable acquisition conditions, and practical adoption would require evaluation across multiple sites and sensors to assess robustness and generalisation.

This limitation is consistent with broader observations in the forestry LiDAR literature. A recent review of AI methods for terrestrial point clouds [183] notes that many deep learning studies rely on very dense TLS data, and that models trained under such favourable conditions may not transfer directly to sparser ALS point clouds unless density effects are explicitly addressed (e.g., through downsampling experiments or multi-density training). At the same time, emerging benchmark studies indicate that deep learning methods can remain competitive across a range of airborne point densities. For example, the FGI-EMIT benchmark for individual tree segmentation from multispectral ALS [184] reports that deep learning approaches outperform traditional unsupervised baselines across multiple density levels. However, segmentation quality degrades as density decreases. Taken together, these findings suggest that extending the present near-ground classification workflow to lower-density airborne or mobile LiDAR is a plausible direction, but it would require systematic evaluation across varying point densities, forest types, and sensor config-

urations before conclusions about operational applicability can be drawn.

A second limitation concerns the size and representativeness of the labelled dataset. Training and evaluation were performed on six manually annotated blocks, with only two blocks held out for testing. Although the block-level split reduces spatial leakage, the limited number of independent test blocks restricts conclusions about variability across terrain conditions. It increases sensitivity to the particular train–test partition. Expanding the number and diversity of labelled blocks would be necessary to better characterise performance variability and robustness across heterogeneous near-ground conditions.

Manual annotation introduces an additional source of uncertainty, especially near the ground where class boundaries in TLS data are inherently fuzzy. In these zones, low vegetation often partially occludes rocks, and many points sit at class interfaces with mixed returns, so a single ”correct” label is not always well-defined. As a result, some apparent misclassification in point-wise error maps may reflect uncertain or inconsistent reference labels rather than genuine model failure. This problem is not unique to the present dataset and has been recognised in recent TLS annotation workflows [175, 185]. Beyond forestry TLS, work on implicit label noise in semantic segmentation [186] similarly shows that even carefully curated datasets can contain hidden noise arising from ambiguous boundaries, occlusion, and annotator variability.

Class imbalance is also relevant for interpreting performance metrics. Because class proportions vary across blocks and rock often dominates the point count, weighted metrics may be disproportionately influenced by the majority classes. It can mask poorer performance on minority or boundary-dominated classes, such as low vegetation. Reporting per-class performance, as done here, is therefore essential, and future evaluations could incorporate boundary-focused analyses to better characterise errors at class interfaces.

Additionally, although data augmentation increased the volume of training samples, it did not increase true environmental diversity. Rotations, jittering, and point dropout can improve robustness to sampling artefacts and partial occlusion, but they do not substitute for additional sites, vegetation types, terrain forms, or rock morphologies. As a result, observed performance may still be dominated by site-specific structure despite extensive augmentation. A more diverse labelled dataset would therefore be required to support model development aimed at broader operational deployment and to evaluate transferability across forest conditions.

Several practical deployment aspects also remain untested. Evaluation was performed on pre-defined blocks rather than on continuous full-scene inference. Operational mapping would require applying the model over large areas using tiling and overlap strategies, which can introduce edge effects and spatial inconsistencies at tile boundaries. Demonstrating full-scene prediction and post-processing for spatial consistency would therefore be necessary to assess practical mapping behaviour beyond block-level metrics. In addition, computational performance was not evaluated. Operational deployment would require assessing runtime, memory requirements, and scalability to larger scenes or near-

real-time data streams from mobile platforms.

Finally, although point-wise metrics indicate semantic separability, they do not directly translate into obstacle-detection performance at the object level. Operational use would require converting predicted labels into discrete obstacle objects (e.g., through polygonisation) and validating detection rates, size estimates, and positional accuracy against independent field measurements. Without object-level evaluation, the relationship between segmentation metrics and trafficability-relevant obstacle metrics remains unresolved.

Taken together, these limitations indicate that the present work should be interpreted as experimental and exploratory rather than as an operational solution for forest applications. The workflow demonstrates that near-ground semantic separation from high-density TLS point clouds is feasible under favourable conditions, but the practical usefulness of the approach for routine forest operations remains untested. In particular, questions related to generalisation across sites and sensors, robustness under lower point densities, full-scene mapping behaviour, object-level obstacle detection, and integration with operational data streams have not yet been addressed.

Nevertheless, the results provide a proof of concept by illustrating how rich geometric and radiometric (intensity) information contained in point clouds can be exploited to characterise near-ground terrain structure. As more diverse labelled datasets become available and as feature representations and validation strategies mature, similar workflows could be adapted to more operational settings, such as machine-mounted or vehicle-based LiDAR systems used in forest harvesting and planning. In this sense, the chapter provides methodological insight rather than a deployable tool. It helps clarify both the potential and current limitations of point-cloud-based approaches for supporting forest operations.

4.6 Conclusion

This chapter demonstrated the feasibility of supervised semantic segmentation for separating three near-ground terrain classes - rock, low vegetation, and ground, from high-density TLS point clouds, achieving encouraging performance on spatially independent test blocks. Including scanner-recorded intensity provided a consistent, though moderate, improvement across aggregate metrics, indicating that intensity can complement geometric cues in complex near-ground scenes. However, low vegetation remained the most challenging class due to occlusion and diffuse boundaries, and the substantial overlap in class-wise intensity distributions limits the use of simple intensity thresholds.

The results should be interpreted as an experimental proof of concept based on a single high-density TLS dataset. Future work should (a) evaluate generalisation across multiple sites, sensors, and point densities, (b) demonstrate full-scene inference with strategies to reduce tiling artefacts, and (c) move beyond point-wise metrics by extracting and validating obstacle objects against independent field measurements to assess operational relevance for trafficability assessment.

4.7 Data availability

The corresponding paper is currently in draft form and will be submitted to a scientific journal in April 2026.

Chapter 5

Terrain complexity assessment in mountain forests using multi-sensor data

5.1 Abstract

Accurate digital terrain models (DTMs) are fundamental for terrain analysis in steep forest environments, yet their reliability depends strongly on sensor type, ground-point density, and spatial resolution. This study evaluates how DTMs derived from airborne laser scanning (ALS), unmanned aerial vehicle (UAV)–borne light detection and ranging (LiDAR), handheld LiDAR (HLS), and UAV photogrammetry (SfM) perform in a steep, closed-canopy mountain forest. Full-density UAV LiDAR served as the internal reference. At plot scale (0.5–4 m grids), vertical deviations were quantified and related to terrain and canopy characteristics using slope, raster-based roughness, topographic wetness index (TWI), and a composite terrain complexity index (TCI) used to stratify terrain into three variability classes. At stand scale (0.5–30 m grids), the influence of sensor choice, LiDAR ground-point density, and grid resolution on slope statistics, slope-category aggregation, and terrain-variability metrics was assessed. HLS and UAV LiDAR provided the most consistent terrain representation under canopy, ALS showed moderate performance, and UAV SfM exhibited large canopy- and terrain-dependent deviations. Terrain roughness was the strongest predictor of elevation error across sensors. Coarser grid resolutions progressively smoothed steep terrain, substantially reducing the mapped extent of high-slope categories. LiDAR DTMs gridded at 2–4 m with approximately 5–10 ground points m^{-2} achieved ~90% agreement with the reference while limiting data volume. Overall, the results demonstrate how sensor characteristics and spatial resolution influence raster-derived terrain indicators and provide methodological guidance for selecting DTM specifications for terrain analysis in complex forest environments.

5.2 Introduction

Mountain forests cover over 9 million km^2 of the Earth's surface, representing about 23% of global forest cover [187]. They provide essential ecosystem services, including carbon

sequestration, biodiversity conservation, and watershed protection. Within this context, European mountain forests cover approximately 41% of the continent's mountain area and serve as a vital timber resource, particularly in Central and Southern Europe [187, 188]. Yet steep and rocky terrain makes mechanised harvesting both technically demanding and environmentally sensitive, and forest operations in these landscapes depend heavily on detailed terrain information to select appropriate machinery, design safe extraction routes, minimise soil disturbance, and manage costs and productivity [21, 23, 22, 24].

The relationship between terrain steepness and machine operability is well established. Conventional cut-to-length (CTL) harvesters and forwarders are generally constrained below about 30–40% slope by traction and stability limits [12], and safety guidance commonly applies similar thresholds when manufacturer limits are unknown [189]. Empirical evidence shows, however, that machines may exceed these nominal limits under specific configurations: wheeled machines with traction aids can reach ~45%, self-levelling tracked systems 60%, and winch-assist systems 75–85% [11]. In very steep terrain, harvesting often shifts to manual felling and cable-based extraction, with cable yarding frequently applied on slopes exceeding 50% [13]. Reliable, spatially explicit slope information is therefore essential for selecting extraction systems and planning operations in mountain forests.

Slope information is increasingly derived from remote sensing, which provides spatially continuous terrain surfaces over large areas more efficiently than field-based surveying [190, 191]. Remote sensing data are typically processed into gridded elevation products: digital surface models (DSMs) represent the uppermost surface, including canopy and built structures, while digital terrain models (DTMs) describe the bare-earth surface beneath vegetation. In image-based workflows, orthomosaics provide complementary visual context for terrain interpretation.

However, DTMs are only as reliable as the underlying ground observations and interpolation methods used to produce them. In steep, closed-canopy mountain forests, vegetation limits ground visibility, steep slopes amplify minor elevation errors into larger slope errors, and terrain complexity introduces microtopographic variability that is easily smoothed at coarser grid resolutions. As a result, DTMs from different sensors and processing choices can differ substantially, and these differences propagate into derived terrain metrics used in operational planning.

A range of remote sensing platforms is available for terrain mapping in forest operations, each entailing distinct trade-offs in coverage, point density, and performance under canopy [49, 52, 47]. Airborne laser scanning (ALS) provides efficient large-area coverage and is widely used as a planning baseline, supporting the derivation of DTMs and susceptibility layers such as depth-to-water (DTW) and topographic wetness index (TWI) for access and trafficability planning [55, 54, 57, 35, 56]. Its representation of fine-scale terrain structure depends strongly on ground-return density and interpolation, with typical national survey specifications of roughly 1–4 points m^{-2} [47, 58].

UAV LiDAR can deliver denser point clouds over smaller areas and better capture complex terrain, making it well-suited for operational planning that requires greater detail. It can be deployed on demand and repeated as needed to reflect current site conditions, making it more flexible than scheduled ALS campaigns for stand-level planning [192]. Compared with typical ALS specifications, UAV LiDAR usually provides substantially higher point densities, enabling more detailed representation of microtopography and small terrain features relevant to extraction planning [59, 60].

Handheld and mobile LiDAR (HLS and MLS) can provide very dense local mapping under canopies, where other sensors fail to capture sufficient ground returns, and GNSS is unreliable. However, these systems cover only limited areas and are susceptible to registration drift. Typically relying on SLAM (simultaneous localisation and mapping) to register scans in GNSS-denied forests [193, 65, 64], they require sensible trajectories and loop closures to keep drift low; when these conditions are not met, accumulated errors can distort terrain geometry and reduce the reliability of derived metrics for operational planning [194].

UAV photogrammetry using structure-from-motion (SfM) can perform well where the ground is visible, but in closed-canopy forests, it often reconstructs canopy surfaces or infers incomplete ground geometry, limiting its suitability for DTM generation [52, 50, 53]. In open or post-harvest conditions, however, passive UAV imaging can be effective for disturbance mapping: orthomosaics and surface models derived from SfM have been used to delineate soil disturbance extent, reconstruct skid-trail deformation, and estimate rut depths from elevation differences [70, 73, 74, 75]. Across all these applications, performance is strongly conditioned by canopy openness, understorey obstruction, and image geometry.

Beyond slope, terrain assessment in forest planning often considers surface ruggedness to describe spatial variability in terrain form at scales relevant for route design. In gridded DTMs, this is commonly quantified using the terrain ruggedness index (TRI), which summarises local elevation variability within a neighbourhood [144], and raster roughness metrics that compute the maximum elevation difference between a central pixel and its surrounding neighbours [195]. Both metrics represent statistical variation in the interpolated terrain surface and describe geomorphometric variability rather than the physical presence of ground obstacles in terms of traditional terrain classification systems [43], so their operational interpretation depends on spatial scale, data resolution, and planning context.

These metrics are inherently scale-dependent and may partly reflect slope trends rather than slope-independent irregularity. Recent work has highlighted the need to interpret TRI-type metrics with caution on inclined terrain and, where appropriate, to apply detrending to isolate slope-independent variability better [79, 84]. These considerations are especially relevant in mountain forests, where steepness and microtopography interact and where sensor-dependent DTM characteristics propagate into both slope and rugged-

ness descriptors used in forest planning.

Despite the widespread use of remote-sensing-derived DTMs in forest planning, there remains limited empirical guidance on how sensor choice, ground-point density, and grid resolution influence the reliability and comparability of raster-derived terrain indicators in steep, closed-canopy mountain forests. While operational planning at tactical and strategic scales often relies on aggregated terrain summaries (e.g., mean slope or dominant terrain class) [43], these indicators are mathematically derived from underlying spatially explicit raster data. Consequently, their stability depends on how terrain variability and extreme values are represented, or smoothed, across spatial resolutions.

The aim of this study is therefore to evaluate how sensor characteristics and spatial resolution influence the reliability and aggregation behaviour of raster-derived slope and terrain-variability metrics in forest-planning contexts. Specifically, the study examines (a) the consistency of elevation surfaces and derived terrain metrics across ALS, UAV LiDAR, HLS, and UAV SfM DTMs, (b) the sensitivity of slope and roughness metrics to grid resolution and LiDAR ground-point density, and (c) how these differences propagate into aggregated slope-based categories used to summarise steepness conditions. By analysing the full data pathway from sensor acquisition to metric aggregation, this work provides methodological guidance for interpreting remote-sensing-derived terrain products in complex forest terrain.

5.3 Materials and methods

This section describes the multi-sensor datasets and processing steps used to generate DTMs and terrain metrics, evaluate DTM vertical accuracy, and assess how sensor type, ground-point density, and grid resolution affect slope-based terrain mapping in steep, closed-canopy forest terrain. An overview of the workflow is provided in Fig. 5.1.

5.3.1 Study area

The study was conducted in a mountain forest region near Cesuna, located on the Asiago Plateau in the Veneto region of northeastern Italy (Fig. 5.2). The area is characterised by steep and complex topography, with elevations ranging from approximately 1050 m to 1300 m above sea level. Slopes vary considerably across the landscape, typically ranging between 15% and 30%, with localised areas exceeding 35%. The terrain includes rocky, uneven ground.

Forest cover is dominated by European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*), forming a mostly closed canopy. The understorey is dense, and ground visibility is limited. Previous forest operations in this area were carried out in less-steep sections using conventional CTL equipment, whereas steeper zones required winch-assisted harvesting.

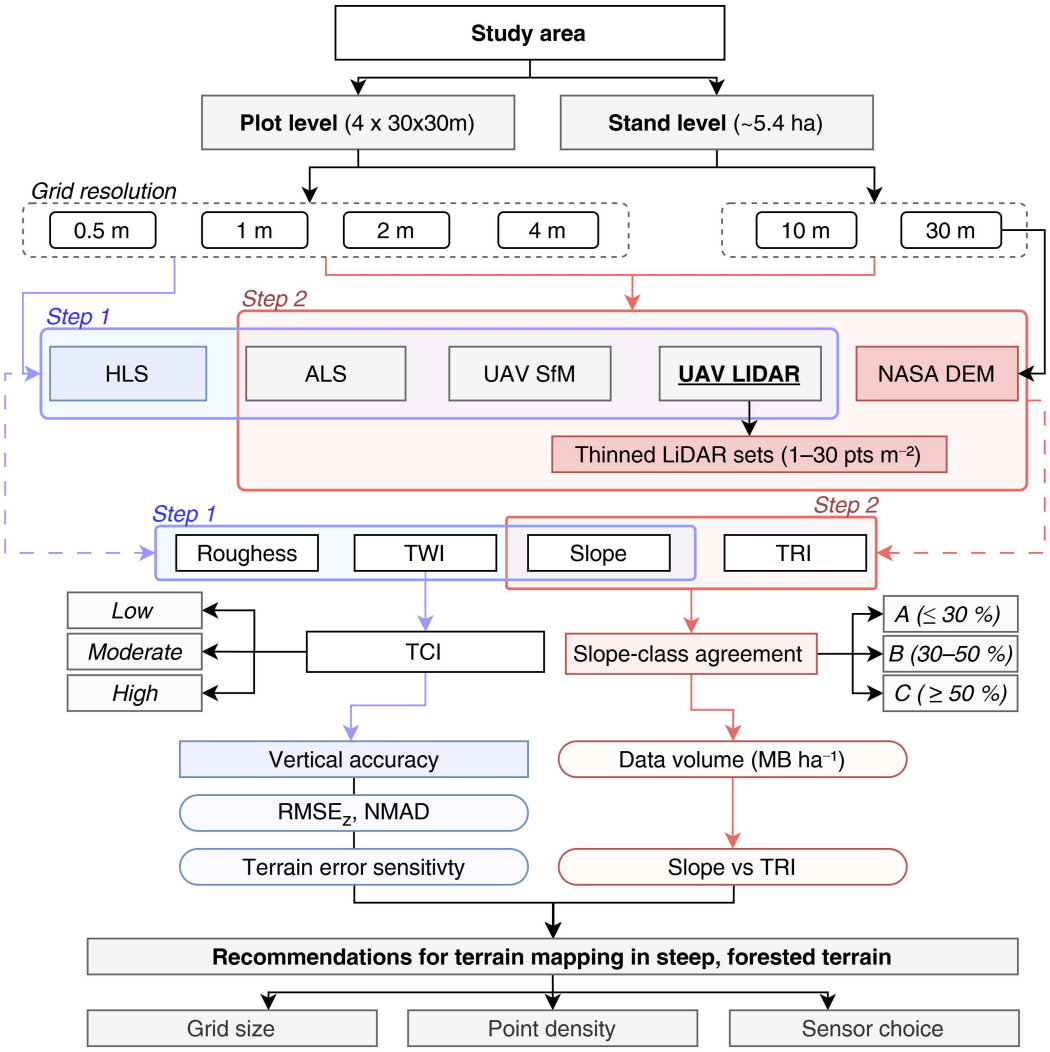


Figure 5.1: Overview of the multi-sensor workflow used to derive terrain metrics, evaluate DTM accuracy, and develop recommendations for terrain mapping in steep, forested terrain. UAV LiDAR (underlined) is used as the reference DTM for all accuracy and slope-class analyses.

5.3.2 DATA ACQUISITION AND PRE-PROCESSING

Four remote sensing techniques - ALS, UAV LiDAR, UAV SfM photogrammetry, and handheld LiDAR, were applied to characterise terrain and its complexity within the study area. The following subsections describe the acquisition details and characteristics of each dataset.

5.3.2.1 ALS DATA

ALS data in the Cesuna area was acquired in 2021 by Terna Driving Energy S.p.A. and subsequently shared with the Veneto Region (Regione del Veneto). According to Terna's public reports [196, 197], the acquisition involved helicopter-based LiDAR inspections using RIEGL VUX-1LR and/or VUX-120 sensors or equivalent instruments. The mission configuration indicates a wide-area "blanket" survey, typically conducted at high altitudes (several hundred metres above ground level, AGL) in a systematic grid pattern, consistent with

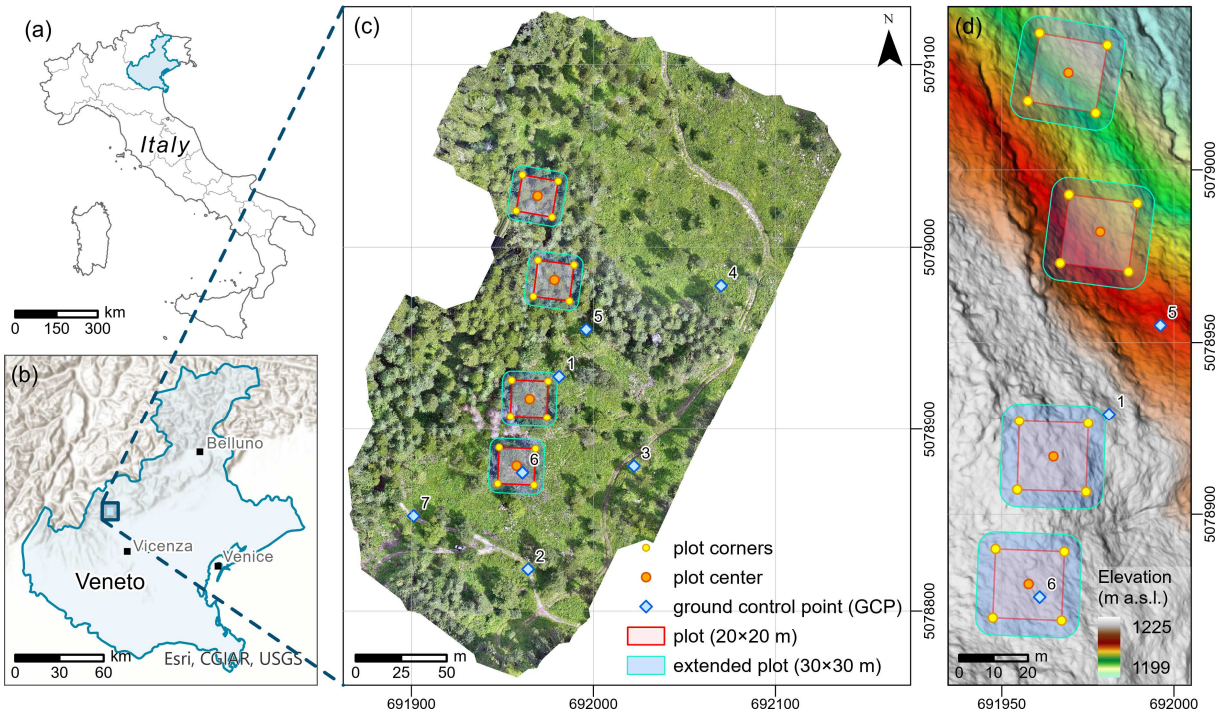


Figure 5.2: Study area overview. (a) Location of the study site in northeastern Italy. (b) Regional context within the Veneto region, showing Cesuna on the Asiago Plateau. (c) UAV-derived orthomosaic with the locations of four 30×30 m handheld LiDAR plots, including plot corners (yellow), centres (orange), and ground control points (GCPs, blue diamonds). (d) DTM-derived elevation map of the same area, illustrating the steep terrain across plot locations.

regional-scale blanket survey practices.

The delivered data set consisted of LAZ tiles georeferenced in EPSG:32632 (WGS 84 / UTM zone 32N), with a nominal point density of approximately 1 to 4 points m^{-2}. The data were provided as classified point clouds. The classification followed the standard ASPRS scheme, distinguishing ground, low- and high-vegetation, and building classes [198]. Detailed flight parameters, such as precise altitude and overlap, were not documented in the final technical report.

5.3.2.2 UAV LiDAR DATA

UAV LiDAR data was collected in June 2025 using a DJI Matrice 300 RTK UAV equipped with a DJI Zenmuse L2 sensor. The survey was carried out at an average altitude of 55 m and a speed of 4.8 metres per second (m/s), covering approximately 5.4 hectares (ha). The mission was planned and initiated on-site using the DJI Pilot 2 app. Data were georeferenced in real-time via the UAV's onboard RTK GNSS, which received correction signals from an Emlid Reach RS2+ base station. The base accessed a regional GNSS correction service to improve absolute positioning accuracy. In addition, seven ground control points (GCPs) were distributed across the survey area using black-and-white ~70×70 cm panel markers (Fig. 5.2). Their positions were measured with an Emlid Reach M2 rover connected to the same RS2+ base station, with points recorded via the Emlid Flow app.

The raw data were recorded in DJI proprietary formats (.LDR, .LDRT, .RPOS, .MRK, .RTK, .RPT) and subsequently processed in DJI Terra (v4.5.0) [199], where the GCPs were used as accuracy and control check points to obtain a classified LAS point cloud, a digital elevation model (DEM), a digital surface model (DSM), and an orthomosaic.

To enable analyses of how ground-point density affects terrain-mapping performance, the full-density UAV LiDAR point cloud was systematically thinned using the `FUSION/LDV ThinData` function [160]. Five downsampled datasets were generated at target ground-point densities of 1, 5, 10, 20 and 30 pulses m^{-2}. Thinning was performed with a 25 m^2 cell size (5×5 m), following `FUSION/LDV` recommendations, and multiple random seeds (option `/rseed`) were used to ensure that independent subsets of points were maintained at each density level. Each run produced a thinned LAS file with the specified target density.

5.3.2.3 UAV SfM photogrammetry data

RGB imagery was collected during the same flight as the UAV LiDAR data, covering approximately 5.4 ha. In this case, only RGB images were used for photogrammetric processing. The images were collected using the DJI Zenmuse L2 RGB camera module, mounted on a DJI Matrice 300 RTK drone. The camera features a 20 MP 4/3″ CMOS sensor with a global mechanical shutter, allowing distortion-free image capture during motion. The images were acquired at approximately 2.3 cm/px and saved as high-quality JPGs with embedded GPS metadata.

All images were processed with Agisoft Metashape Professional (v2.2.0) [122], using the same GCPs used in the UAV LiDAR workflow to ensure consistent georeferencing. The photo alignment was performed with high accuracy, using a keypoint limit of 150,000 and with adaptive camera model fitting enabled. The dense point cloud was generated from depth maps using high-quality and moderate depth filtering. The resulting photogrammetric output included a dense point cloud, a high-resolution orthomosaic, and a DEM. The point cloud derived from this processing was used in further terrain analysis.

5.3.2.4 Handheld LiDAR data

Handheld LiDAR scanning (HLS) data was acquired using the STONEX X120GO scanner, a mobile 3D mapping device capable of real-time data acquisition and post-processing via the GOapp/GOpost software [200]. The system operates with 16 laser channels and has an approximate field of view of 360° × 270°. It achieves a vertical resolution of up to 6 mm in controlled environments, scanning at 320,000 points per second.

Four rectangular forest plots (20 × 20 m) were surveyed by walking a series of parallel, back-and-forth transects at a constant pace, ensuring complete ground coverage within each plot and a buffer zone around each plot. Due to the scan path and the buffer margin, the effective area per plot was approximately 30 × 30 m (∼ 0.16 ha). For plot layout, the plot centre was first established in the field, and then the four corners of the 20 × 20 m square were determined using compass bearings and distance measurements (10 m from

the centre in cardinal directions). To ensure visibility of the reference corner points in the LiDAR scans, expanded polystyrene spheres (white, ~20 cm diameter) were mounted on magnetised 1 m poles placed at the four corners (Fig. 5.3). The positions of these corner markers were measured with an Emlid Reach M2 rover. Each scan lasted 13–15 min at a slow walking pace (≈ 1.0-1.3 m s^{-1}), with total point densities of 50–70 × 10^3 points m^{-2}, estimated as total points divided by the 2-D convex-hull footprint of the point cloud.

Figure 5.3: Examples of ground control and reference targets used in the study. (a) UAV ground control marker (~70×70 cm black–white panel) used for georeferencing UAV LiDAR and photogrammetry; (b) expanded polystyrene sphere (~20 cm diameter) mounted on a magnetised 1 m pole, placed at the corners of HLS plots for GNSS positioning; (c) corner-sphere target as captured in the handheld LiDAR point cloud.

During acquisition, HLS point clouds were recorded in a local frame. Georeferencing to EPSG:32632 (UTM zone 32N) was done in *CloudCompare* using GNSS-measured corner markers as control points and the *Align tool (point-pairs picking)* with at least four correspondences. The precision of the control points varied by plot, particularly under canopy (Table 5.7). To ensure consistent inter-dataset registration, each HLS cloud was additionally aligned with the UAV LiDAR reference as described in Section 5.3.3.1 through the same *Align tool (point-pairs picking)* tool followed by the iterative registration of the closest point algorithm (ICP) .

5.3.3 General processing steps

An overview of the multi-sensor workflow is provided in Fig. 5.1. All processing was organised at two spatial scales. At the plot level, analyses were restricted to the four 30 m × 30 m reference plots (Fig. 5.2) where all high-resolution datasets overlap (ALS, UAV LiDAR, UAV SfM and HLS). At the stand level (the full ~5.4 ha UAV LiDAR footprint), coarser grid sizes of 10 and 30 m were also considered to assess how resolution affects slope-based terrain categorisation and terrain ruggedness patterns (Fig. 5.1). The following subsections describe each processing step in detail.

5.3.3.1 Cross-dataset alignment

To minimise residual offsets and ensure consistent spatial registration prior to accuracy and terrain-metric comparisons, we adopted the UAV–LiDAR point cloud as the reference dataset. ALS, UAV SfM, and HLS products were rigidly aligned to the UAV–LiDAR data in *CloudCompare*. We first applied the *Align (point pairs picking)* tool using at least seven well-distributed point pairs, then refined the transformation using iterative closest point (ICP) while fixing the scale to 1.0. For each registration, we recorded *CloudCompare's* quality outputs, including the ICP final RMS (root-mean-square of residual distances after alignment) and the scale setting. These RMS values quantify the relative registration error between the aligned dataset and the UAV-LiDAR reference.

5.3.3.2 Ground classification

After alignment, all point clouds (ALS, HLS, UAV LiDAR, UAV SfM, and all thinned UAV LiDAR datasets) were classified with LAStools `lasground64` using `-wilderness` (step=3), `-ultra_fine` (8 substeps for steep terrain), `-all_returns`, and `-compute_height` [159].

The outputs were normalised to the ASPRS classes (ground = 2). For denser HLS plots, an additional pass with CloudCompare's cloth simulation filter (CSF) [**?**] was applied (*Scene* = steep slope, *cloth_resolution* = 0.95 m, *max iterations* = 1000, *classification threshold* = 0.025) to remove residual vegetation points and improve ground classification. The parameters were selected through trial and error and vary with the area's terrain and canopy conditions.

Furthermore, to characterise canopy cover and ground exposure across plots, we processed non-ground (canopy) and ground full–density UAV–LiDAR returns with `FUSION/LDV ReturnDensity` (default settings) on a 0.5 m grid. The tool outputs a raster of return density at the specified resolution and a binary mask layer indicating canopy presence, with pixels coded as 1 for canopy and 0 for no canopy. In this case, the UAV LiDAR dataset was used, as it reflects the latest canopy conditions and provides robust spatial positioning via RTK GNSS and GCPs.

Void proportion, defined as the percentage of pixels with zero ground returns, indicates areas where the laser failed to penetrate the canopy to the ground. Spatial clustering of void pixels was quantified using Moran's I statistic [201], a global spatial autocorrelation index ranging from -1 (dispersed) to $+1$ (clustered). Values near $+1$ indicate clustered voids typical of dense canopy conditions.

5.3.3.3 DTM generation

The classified ground points were converted to raster DTMs using the FUSION/LDV `GridSurfaceCreate` tool. `GridSurfaceCreate` constructs a gridded surface by averaging the elevations of points within each cell and interpolating across empty cells [160]. No median or mean smoothing filters were applied to preserve fine-scale detail.

Overlap points were excluded using the `/ignoreoverlap` parameter. Raster hole filling was restricted with conservative radii to limit spatial smoothing (units are cells; default is 99). These settings reduced artefacts in thinned datasets while still avoiding small gaps. Specific fill-distance parameters were selected through trial and error and adjusted based on grid resolution and data density.

In addition to the point cloud–derived DTMs, we incorporated the publicly available NASA DTM raster product covering the study area [202]. The corresponding DTM tile was reprojected to UTM zone 32N and aligned to the UAV LiDAR reference grid using nearest-neighbour resampling. The NASA DTM was used at its native 30 m resolution for the slope-class comparison, without any additional downscaling or smoothing.

5.3.3.4 Terrain metric computation

All DTMs were aligned to the UAV LiDAR reference grid to ensure identical pixel origin, extent, and CRS. Raster alignment and clipping were performed in a QGIS Python environment combining native GDAL, SAGA GIS, and QGIS algorithms. A custom function handled reprojecting, snapping to the reference grid, and ensuring consistent spatial registration.

For plot-level analysis, all DTMs at resolutions of 0.5, 1, 2, and 4 m were clipped to the 30×30 m plot boundaries (Fig. 5.2), with plot boundaries buffered by 1 pixel to avoid edge artefacts. For stand-level analysis, terrain metrics were computed on the full ~5.4 ha UAV LiDAR footprint without clipping, allowing examination of broader spatial patterns across the entire surveyed area.

Four terrain metrics that are widely used in forestry and forest operations studies [195, 152, 203, 204, 24] were considered: slope, surface roughness, TWI and terrain ruggedness index (TRI). At the plot level, slope, roughness and TWI were computed to characterise surface complexity. At the stand level, slope and TRI were used to examine broader spatial patterns of terrain classification and ruggedness. The computational definitions and parameters for all metrics follow [205, 195, 55, 206, 207] and are detailed in Table 5.1. All metrics were derived with GDAL and SAGA GIS.

The output of this workflow consisted of (a) plot-level rasters of slope, roughness and TWI for all high-resolution sensors (ALS, UAV LiDAR, thinned UAV LiDAR, UAV SfM and HLS) at 0.5, 1, 2 and 4 m, and (b) stand-level rasters of slope for ALS, UAV SfM, full-density and thinned UAV LiDAR and the NASA DTM at 0.5, 1, 2, 4, 10 and 30 m, together with stand-level TRI rasters for the full-density UAV LiDAR at the same resolutions (Fig. 5.1). For statistical analyses, raster values were exported as tabular data, with each pixel represented by a row and each metric by a column, grouped by sensor, spatial scale (plot or stand) and resolution. These tabular datasets were then used to compute pixel-wise accuracy statistics for each metric and sensor, to derive the TCI, and to analyse slope/TRI relationships across resolutions.

Table 5.1: Raster-based terrain metrics applied to sensor-derived DTMs. At the plot level, metrics were computed for ALS, UAV LiDAR (full and thinned), UAV SfM and HLS at grid resolutions of 0.5, 1, 2 and 4 m. At the stand level, slope was derived for all DTMs covering the full survey area (ALS, UAV LiDAR, thinned UAV LiDAR, UAV SfM and the NASA DTM), and TRI was computed from the full-density UAV LiDAR DTM at 0.5, 1, 2, 4, 10 and 30 m.

Metric	Description	Ref.	Units	Res.	Parameters	Software
Slope	Gradient magnitude of DTM. Low: gentle; high: steep terrain	[205]	° (deg) → %	0.5, 1, 2, 4, 10, 30 m	Per pixel	SAGA GIS
TRI (raw)	Terrain ruggedness index: SD of elevation in a moving window; combines steepness and micro-relief	[195]	m	0.5, 1, 2, 4, 10, 30 m	Square window with radii 2, 4, 6 m (three variants)	GDAL, SAGA GIS
Roughness	Local SD of elevation in 5×5 m window. Low: smooth; high: rough/rocky	[195]	m*	0.5, 1, 2, 4 m	Window: 5×5 m	GDAL
TWI	Topographic wetness index. Low: dry ridges; high: wet valleys	[55, 206, 207]	—**	0.5, 1, 2, 4 m	Flow dir: D8	SAGA GIS

* Roughness units are based on the input DTM. Since the DTM is in metres, roughness is expressed in metres (m).
** TWI is a unitless index representing relative wetness potential computed from upslope contributing area and local slope.

5.3.4 Statistical analysis

5.3.4.1 Terrain complexity evaluation

To quantify and stratify terrain complexity within the four study plots (Fig. 5.2) for subsequent analyses, we developed a simple terrain complexity index (TCI) based on terrain metrics calculated from the data. Four external datasets were used, representing diverse terrain conditions across multiple study regions (Table 5.2), to determine the weights of the objective metric. For each sensor and spatial resolution, terrain metrics were derived at the pixel scale, providing comprehensive sets of slope, roughness, and TWI values. All metrics for these datasets were computed following the same processing pipeline described in Sections 5.3.3.2, 5.3.3.3, and 5.3.4.1.

To ensure balanced and representative sampling across all terrain types, from flat to steep and smooth to rugged, a stratified sampling approach was applied. Within each external dataset, pixels were first sorted by slope values and divided into quantile bins, effectively grouping pixels from very flat to very steep. An equal number of pixels were sampled from each bin within each external study area, and terrain metrics were extracted at these

Table 5.2: Description of reference study sites, data collection sensors, landscape, and operational context used for weight calculation with the CRITIC algorithm. All datasets were collected by the authors and used solely to estimate metric weights.

Dataset	Location	Sensor	Vegetation	Terrain	Operations
DS1 (Paneveggio)	Paneveggio, Trentino-Alto, Italy	UAV LiDAR	Spruce-dominated forest	Steep, high-relief, rough; highest elevation.	Mechanised harvesting is feasible but constrained by slope
DS2 (Cesuna)	Cesuna, Veneto, Italy	UAV LiDAR	Mixed spruce-beech forest	Very steep, highly rocky; second highest complexity	Mechanized operations possible only on gentler sections; rockiness and slope limits mobility
DS3 (Jonkershoek)	Jonkershoek Nature Reserve, Western Cape, South Africa	TLS	Pine plantations; area affected by recent fire with sparse/no cover	Very rocky surface; moderate slopes	Historically motor-manual felling; cable skidder extraction typical
DS4 (Sabie)	Sabie, Mpumalanga, South Africa	UAV LiDAR	Pine plantations	Gentlest and smoothest; wetness highest	Conventional mechanized harvesting and extraction widely feasible

locations. This method prevented any single region or extreme terrain type from dominating the analysis.

The sampled pixel data were normalised using robust min-max scaling (2nd to 98th percentiles) to exclude extreme outliers. Subsequently, the objective metric weights were derived using the CRITIC (criterion importance through intercriteria correlation) algorithm. The CRITIC method assigns weights to criteria based on two components: the standard deviation (contrast strength) of each metric and its correlation independence with other metrics. It was selected because it objectively derives weights based on both the variance (contrast strength) and the independence of each metric. Formally, the weight W_k for the metric k is calculated as:

$$W_k = \frac{C_k}{\sum_{j=1}^{m} C_j} \tag{5.1}$$

where $C_k = \sigma_k \sum_{i=1}^{m}(1 - r_{ik})$ is the contrast measure for the metric k, σ_k is the standard deviation of the metric k, r_{ik} is the Pearson's correlation coefficient between the metric i and k, and m is the total number of metrics. Metrics with high variability and low cor-

relation with other metrics receive higher weights. The weighted combination of slope, roughness, and TWI was used to compute the TCI for each pixel. The TCI was calculated according to Eq. 5.2.

$$\mathrm{TCI}(\mathrm{px}) = w_s \times S_{\mathrm{norm}}(\mathrm{px}) + w_r \times R_{\mathrm{norm}}(\mathrm{px}) + w_t \times \mathrm{TWI}_{\mathrm{norm}}(\mathrm{px}) \tag{5.2}$$

where w_s, w_r, and w_t are the CRITIC-derived weights of Eq. 5.1, and S_{norm}, R_{norm}, and $\mathrm{TWI}_{\mathrm{norm}}$ are the normalised slope, roughness and TWI values, respectively, scaled from 0 to 1 using min-max scaling (2nd to 98th percentiles). Since all metrics were normalised, the resulting TCI ranges from 0 (lowest complexity) to 1 (highest complexity). This weighted index was then applied to the four study plots.

TCI classes were defined using global tertiles of the UAV LiDAR–TCI as the overall reference, with the distribution pooled across all plots at each resolution. For each resolution r, the 33rd and 67th percentiles of TCI, denoted $P_{33}^{(r)}$ and $P_{67}^{(r)}$, were computed from all pixels at that resolution and applied as fixed thresholds to every plot to define low complexity $\{\mathrm{TCI} \leq P_{33}^{(r)}\}$, moderate complexity $\{P_{33}^{(r)} < \mathrm{TCI} \leq P_{67}^{(r)}\}$, and high complexity $\{\mathrm{TCI} > P_{67}^{(r)}\}$ classes (Table 5.3).

Table 5.3: Interpretation of terrain complexity classes derived from UAV LiDAR. Classes are defined by global tertiles (P_{33}, P_{67}) of the TCI across all plots and resolutions, and are summarised in terms of typical terrain conditions and indicative operational implications based on slope, surface roughness, and TWI, under the assumption that increasing terrain complexity generally requires progressively more specialised harvesting and extraction systems.

Class	TCI range	Terrain summary	Indicative operational requirements
Low	$\mathrm{TCI} \leq P_{33}$	Gentle, smooth, and easily accessible terrain	Standard CTL systems; conventional ground operations
Moderate	$P_{33} < \mathrm{TCI} \leq P_{67}$	Mixed conditions with intermediate accessibility and variable terrain features	Cable-assist systems or specialized equipment recommended
High	$\mathrm{TCI} > P_{67}$	Steep and irregular topography requiring specific handling and safety considerations	Winch or cable-based extraction systems required

5.3.4.2 Vertical and metric accuracy

To first evaluate vertical accuracy at the plot level, the DTMs of ALS, HLS, and UAV SfM were compared against the UAV LiDAR reference for each plot and spatial resolution (0.5, 1, 2, and 4 m). The accuracy assessment followed the ASPRS guidelines for elevation data [208], using two metrics: root-mean-square error (RMSE, the primary ASPRS-

recommended measure) and normalised median absolute deviation ($NMAD_z$, robust to outliers). For each plot and resolution, elevation differences were calculated at the pixel level as $d_i = z_{i,\text{test}} - z_{i,\text{ref}}$. The accuracy metrics are defined as in Eqs. 5.3 and 5.4.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_{i,\text{test}} - m_{i,\text{ref}})^2}, \tag{5.3}$$

$$\text{NMAD}_z = 1.4826 \cdot \text{med}\left(\left|d_i - \text{med}(d)\right|\right), \tag{5.4}$$

where d_i is the difference in the evaluated metric (e.g., elevation, terrain metric) at pixel i, n is the total number of pixels, $m_{i,\text{test}}$ and $m_{i,\text{ref}}$ are the metric values from the evaluated and reference datasets, respectively. For vertical accuracy, $d_i = z_{i,\text{test}} - z_{i,\text{ref}}$. The notation $\text{med}(x)$ denotes the median value of x.

5.3.4.3 Terrain-dependent error modelling

To examine how local terrain properties and canopy cover influence DTM accuracy at the plot level, we modelled absolute vertical error (E_i) at the pixel scale. For the non-reference sensors (ALS, HLS, UAV SfM), errors were related to terrain metrics (slope, roughness, TWI) and canopy cover, and were computed relative to the full-density UAV LiDAR DTM.

For each sensor (ALS, HLS, UAV SfM) at each resolution (0.5, 1, 2 and 4 m), we computed absolute vertical error at each pixel as described in Eq. 5.5.

$$|E_i| = \left|Z_{\text{sensor},i} - Z_{\text{reference},i}\right| \tag{5.5}$$

where E_i is the absolute vertical error at pixel i (m), $Z_{\text{sensor},i}$ is the elevation from the test sensor at pixel i, and $Z_{\text{reference},i}$ is the elevation from the reference dataset at pixel i. These errors were then regressed against the three terrain metrics extracted from the UAV LiDAR DTM using a moving 5×5 m focal window centred on each grid cell (Fig. 5.4). The window size was selected to generate many pixel-level terrain measurements within each plot, rather than a single averaged value, and to resolve local variation in terrain–error relationships.

$$|E_i| = \beta_0 + \beta_1 \cdot \text{Roughness}_i + \beta_2 \cdot \text{Slope}_i + \beta_3 \cdot \text{TWI}_i + \varepsilon_i, \tag{5.6}$$

where β_0 is the intercept, β_j are partial regression coefficients representing the relationship between terrain metric j and error, and ε_i is the residual.

Separate ordinary least squares (OLS) models were fitted for each sensor–resolution combination, with standard errors clustered at the plot level to account for spatial dependence. In addition to terrain metrics, the local canopy cover fraction was calculated for the same 5×5 m windows by averaging the binary canopy mask (1 = canopy, 0 = open) within each grid cell, yielding pixel-level canopy values for all plots and resolutions. To assess how

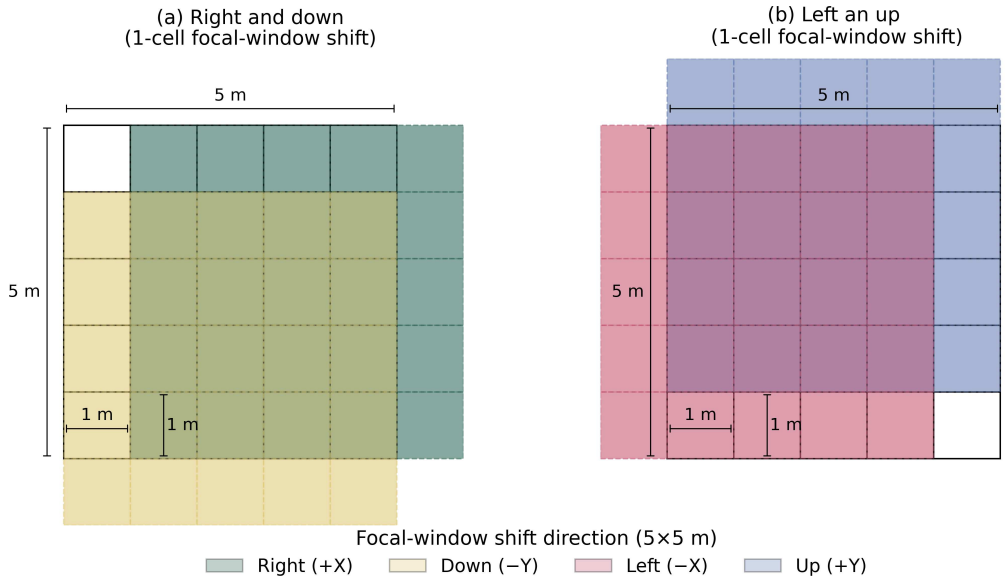


Figure 5.4: Focal moving-window approach for pixel-scale terrain characterisation. A 5 m × 5 m neighbourhood is sequentially centred on each grid cell and shifted across all cardinal directions (panels (a), right/down; panel (b), left/up). At each position, roughness, slope, TWI, and canopy coverage are extracted from the window.

canopy cover relates to DTM error, $|E_i|$ was also regressed and correlated with the local canopy fraction, allowing the influence of tree cover and open ground on DTM accuracy to be distinguished.

Model outputs include coefficient estimates (β_j) quantifying the partial change in absolute error per unit change in terrain metric j, p-values testing the null hypothesis that $\beta_j = 0$ ($\alpha = 0.05$), and the coefficient of determination (R^2), representing the proportion of error variance explained.

5.3.4.4 SLOPE-BASED CATEGORY AGREEMENT

To evaluate how DTM differences propagate into derived terrain information, we quantified agreement in slope-based terrain categories derived from each DTM with the full-density UAV LiDAR reference across the survey area of ~5.4 ha. All slope values (reference and sensors) were first classified into NTC slope classes 1–7 based on the thresholds described in [1]. To link these classes to typical system choices in steep-terrain harvesting, the seven NTC slope classes were then aggregated into three slope-based terrain categories A, B, and C, as described in Table 5.4.

For each sensor-resolution combination, the reference and sensor-slope-based categories (A, B, C) were compared pixel by pixel. Agreement was quantified using three summary metrics: overall agreement (OA), underestimation (UE), and overestimation (OE), as defined in Eqs. 5.7–5.9 and summarised in Table 5.5.

$$\mathrm{OA} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(s_i = r_i) \times 100, \tag{5.7}$$

Table 5.4: Slope-based terrain categories (A, B, C) derived from NTC slope classes [1] and typical machinery operating limits reported in the forest operations literature [209, 189]

Cat.	NTC class	Slope range	Terrain class	Operational interpretation and equipment
A	1–3	≤ 30 %	Gentle to moderate	Wheeled cut-to-length systems generally feasible; conventional ground-based harvesting suitable without assistance
B	4–5	30–50 %	Steep	Winch-assisted or specialised ground-based systems often required; tracked machines or cable-assist tethered equipment
C	6–7	≥ 50 %	Very steep	Cable-based or non-mechanised extraction usually preferred; ground-based harvesting marginal or infeasible

$$\mathrm{UE} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(s_i < r_i) \times 100, \tag{5.8}$$

$$\mathrm{OE} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(s_i > r_i) \times 100, \tag{5.9}$$

where r_i and s_i are the reference and sensor terrain categories at pixel i ($r_i, s_i \in \{A, B, C\}$), n is the total number of pixels, and $\mathbf{1}(\cdot)$ is the indicator function, which returns 1 if the stated condition is true and 0 otherwise.

Table 5.5: Summary metrics for evaluating agreement between reference and sensor-derived slope-based terrain categories. Three primary metrics quantify overall and directional misclassification.

Metric type	Abbr.	Definition and interpretation
Overall agreement	OA (Eq. 5.7)	Percentage of pixels assigned to the same category as the reference DTM. Provides overall classification accuracy across classes A, B and C.
Underestimation	UE (Eq. 5.8)	Percentage of pixels assigned to a lower category than the reference (e.g. reference B mapped as A). Indicates systematic underestimation of slope and associated operational constraints.
Overestimation	OE (Eq. 5.9)	Percentage of pixels assigned to a higher category than the reference (e.g. reference A mapped as B or C). Indicates systematic overestimation of slope and associated operational constraints.

To complement the slope-category agreement analysis with a practical consideration of data handling, total data volumes were quantified for each sensor–resolution combina-

tion included in the stand-level analysis (ALS, UAV SfM, full-density UAV LiDAR, all thinned UAV LiDAR datasets, and the NASA DEM). For each DTM, the total file size (MB) of all associated outputs, including the point cloud (.las), DTM rasters (.asc and .tif), and derived slope raster (.tif), was summed and normalised by the mapped area (ha) to obtain a data-volume metric expressed in MB ha^{-1}. Values were computed consistently across resolutions (0.5, 1, 2, 4, 10, and 30 m) using identical file formats and compression settings.

5.3.4.5 Slope and TRI across scales

Slope is the primary variable used to define the slope-based terrain categories (A, B, C), but it does not distinguish between smooth and highly irregular surfaces. TRI is therefore often used in conjunction with slope to assess trafficability and operational hazards in steep-terrain forestry. To understand how these metrics interact across the range of resolutions relevant for harvest planning, slope and TRI were derived from the full-density UAV LiDAR dataset at grid sizes of 0.5, 1, 2, 4, 10 and 30 m. For each grid size, TRI was computed for three square neighbourhood radii (2, 4 and 6 m), yielding $\mathrm{TRI}_{r,2}$, $\mathrm{TRI}_{r,4}$ and $\mathrm{TRI}_{r,6}$ variants that represent progressively broader terrain context (Table 5.1). In what follows, TRI denotes the original (raw) TRI values before detrending.

For each combination of grid size r and TRI radius R, the Pearson correlation coefficient between slope (%) and TRI was calculated as described in Eq. 5.10.

$$\rho_{r,R} = \mathrm{cor}\left(\mathrm{slope}_r, \mathrm{TRI}_{r,R}\right), \tag{5.10}$$

Statistical significance was tested using two-sided p-values under the null hypothesis of no linear association ($H_0: \rho_{r,R} = 0$). Due to the large number of pixels, almost all correlations were statistically significant ($p < 0.001$). Interpretation therefore focuses on the effect sizes $\rho_{r,R}$ and their change with resolution rather than on p-values. To quantify how variability changes with grid size, we computed the coefficient of variation (CV) for slope and each TRI variant (Eq. 5.11).

To quantify how variability changes with grid size, we computed the coefficient of variation (CV) for slope and each TRI variant (Eq. 5.11).

$$\mathrm{CV}_X = \frac{\sigma_X}{\mu_X}, \tag{5.11}$$

where μ_X is the mean and σ_X is the standard deviation of variable X (either slope or TRI). Because TRI partly reflects slope, we further derived a detrended TRI that isolates ruggedness independent of mean slope. For each resolution and TRI radius, we fitted a simple linear model described in Eq. 5.12.

$$\mathrm{TRI}_{r,R} = \alpha_{r,R} + \beta_{r,R} \cdot \mathrm{slope}_r + \varepsilon_{r,R}, \tag{5.12}$$

where $\alpha_{r,R}$ is the intercept, $\beta_{r,R}$ is the slope coefficient and $\varepsilon_{r,R}$ is the residual error. Detrended TRI was then defined as Eq. 5.13.

$$\mathrm{TRI}^{*}_{r,R} = \varepsilon_{r,R}, \tag{5.13}$$

where $\mathrm{TRI}^{*}_{r,R}$ is the detrended terrain ruggedness index (the residual component). This quantity captures the portion of terrain ruggedness that is independent of slope, i.e. local elevation variability that the background slope at that grid size cannot explain.

Correlation and coefficient-of-variation analyses were then repeated for both raw TRI ($\mathrm{TRI}_{r,R}$) and detrended TRI ($\mathrm{TRI}^{*}_{r,R}$) across all resolutions, allowing comparison of how slope-dependent and slope-independent ruggedness behave at different grid sizes. In addition, to illustrate ruggedness differences within machine-relevant slope categories, raw and detrended TRI were examined within the 30–50% slope band (slope-based class B), and their distributions and relationships with slope were visualised using scatter plots and histograms.

5.4 Results

5.4.1 Terrain complexity classes

The TCI was computed from slope, roughness, and TWI, and the study plots were stratified into three complexity classes: low, moderate, and high. CRITIC-derived weights for the three metrics varied only slightly across four resolutions (Table 5.6). External reference datasets consistently indicated TWI as the dominant component (weights 0.465–0.498), whereas slope and roughness had similar, smaller contributions (0.246–0.273). The relative stability of weights is consistent with a terrain-complexity structure at resolutions of 0.5 m to 4 m. This indicates that observed differences in TCI classes across resolutions reflect scale effects in the terrain metrics rather than changes in metric weighting.

Table 5.6: CRITIC-derived weights for slope, roughness, and TWI at four plot-level spatial resolutions (0.5, 1, 2, and 4 m), based on four external reference datasets. These weights are used to calculate the terrain complexity index (TCI; Eq. 5.2) by combining normalised slope, roughness, and TWI values.

Resolution	Slope	Roughness	TWI
0.5 m	0.257	0.246	0.498
1 m	0.261	0.251	0.488
2 m	0.265	0.256	0.478
4 m	0.263	0.273	0.465

5.4.2 Georeferencing and vertical accuracy

Field GNSS control included seven UAV LiDAR GCPs ($RMSE_x = 0.02$ m, $RMSE_z = 0.01$ m) and sixteen HLS plot-corner markers ($RMSE_x = 0.01$–0.39 m, $RMSE_z = 0.01$–0.52 m), with larger errors in plots 3–4, which were characterised by dense, closed canopy, leading to degraded GNSS visibility and reduced ground returns (Table 5.7). UAV LiDAR georeferencing in *DJI Terra* produced control-point $RMSE_z = 0.02$–0.05 m, checkpoint $RMSE_z = 0.07$–0.09 m, and a mean vertical offset $\Delta Z \approx -0.01$ m (SD 0.02 m). UAV SfM checkpoints had $RMSE_z = 0.37$ m (SD 0.25 m). Post-registration RMS residuals in *CloudCompare* were 0.07–0.09 m for HLS to UAV LiDAR and 0.37 m for ALS to UAV LiDAR, with scale deviation < 0.01 (Table 5.7).

DTM vertical accuracy relative to the UAV LiDAR reference (Fig. 5.5) varied systematically by sensor, canopy condition and terrain complexity. HLS exhibited the lowest and most stable errors across resolutions ($RMSE_z$ 0.18–0.33 m; NMAD 0.07–0.31 m), with little sensitivity to TCI class. ALS performed moderately well, overlapping the upper range of UAV LiDAR in open to moderately closed canopy, but errors increased in dense canopy ($RMSE_z$ 0.54–0.65 m; NMAD 0.31–0.35 m), with slightly higher values in high-complexity terrain (up to 0.55 m in high TCI vs 0.45 m in low TCI). UAV SfM was strongly canopy- and terrain-dependent, with extreme errors in dense plots ($RMSE_z$ up to 6.66 m; NMAD up to 2.77 m) and consistently higher errors than ALS even in open subsets (e.g. 0.91–3.18 m across TCI classes). For HLS and ALS, grid resolution (0.5–4 m) had negligible influence on $RMSE_z$ and NMAD (variation < 0.15 m), whereas for SfM, coarser grids (2–4 m) sometimes reduced extremes in dense plots (e.g. 2.45 m vs. 6.66 m in plot 3) but occasionally amplified errors in open areas, indicating resolution-dependent rasterisation artefacts in sparse point clouds.

5.4.3 Error sensitivity

Pixel-wise analysis relating absolute vertical errors to terrain metrics (roughness, slope, TWI) and canopy cover revealed that local terrain complexity explains a significant portion of the variance in DTM error across sensors and resolutions (18–38%, Table 5.8). Among these metrics, roughness emerged as the strongest and most consistent predictor, with coefficients ranging from +0.02 to +0.16 m of error increase per unit of roughness, especially pronounced for lower-point-density sensors such as ALS and UAV SfM. Slope and TWI effects were generally smaller and less consistent across sensors and resolutions. HLS showed near terrain-independent accuracy, likely due to its ultra-dense point clouds (~14 pts/m^2) providing comprehensive ground coverage, resulting in negligible roughness effects (+0.01–0.04 m) (mostly statistically non-significant). Notably, errors increased by only about 2.1 times from open to dense canopy plots (0.12–0.25 m), suggesting that positioning uncertainty and SLAM drift, rather than terrain complexity, primarily drive error in dense forests for HLS.

Table 5.7: Georeferencing and registration metrics in the UAV LiDAR reference frame. Absolute checkpoint accuracy is reported for UAV LiDAR. SfM figures correspond to absolute photogrammetric checkpoint errors from Agisoft Metashape using independent checkpoints. HLS and ALS entries are relative post-registration residuals to UAV LiDAR (point-pairs preregistration + ICP; 3D RMS, scale deviation in parentheses where applicable). Relative figures are not independent absolute accuracies.

Dataset	Description	Metric	Error (m)
GNSS GCPs (7)*	Field-measured GNSS control accuracy	RMS error - $RMSE_h$ / $RMSE_Z$**	0.02 / 0.01
HLS GNSS GCPs (16)*	Plot-level GNSS accuracy	RMS error - $RMSE_h$ / $RMSE_Z$**	0.01–0.39 / 0.01–0.52
		Plot 1 (4)*	0.01 / 0.01
		Plot 2 (4)*	0.11 / 0.12
		Plot 3 (4)*	0.33 / 0.52
		Plot 4 (4)*	0.34 / 0.24
UAV LiDAR (*DJI Terra*)	**Absolute** checkpoint accuracy (vertical-only)	Control point $RMSE_Z$ (GNSS GCPs 2–5)	0.02–0.05
		Checkpoint $RMSE_Z$ (GNSS GCPs 6–7)	0.07–0.09
		Altitude difference (ΔZ)	–0.05 / +0.03
		Mean vertical offset	~0.01 (SD = 0.02)
UAV SfM (*Agisoft Metashape*)	**Absolute** checkpoint accuracy (independent checkpoints n=3)	Checkpoint $RMSE_h$ / $RMSE_z$	0.37 / 0.25
HLS → UAV LiDAR (*CloudCompare*)	**Relative** registration to UAV frame	Post-registration RMS (3D) to UAV LiDAR (point-pairs + ICP; scale deviation)	0.07–0.08 (<0.01)***
		Plot 1 (4)*	0.07 (<0.01)
		Plot 2 (4)*	0.09 (<0.01)
		Plot 3 (4)*	0.08 (<0.01)
		Plot 4 (4)*	0.08 (<0.01)
ALS → UAV LiDAR (*CloudCompare*)	**Relative** registration to UAV frame	Post-registration RMS (3D) to UAV LiDAR (point-pairs + ICP; scale deviation)	0.37 (<0.01)***

* Number of ground control points measured.
** $RMSE_h$ and $RMSE_Z$ denote horizontal (planimetric) and vertical components of survey accuracy, respectively.
*** Post-registration RMS (3D) = RMS point-to-point distance after point-pairs + ICP (CloudCompare), combining X–Y–Z. Parentheses report scale deviation.

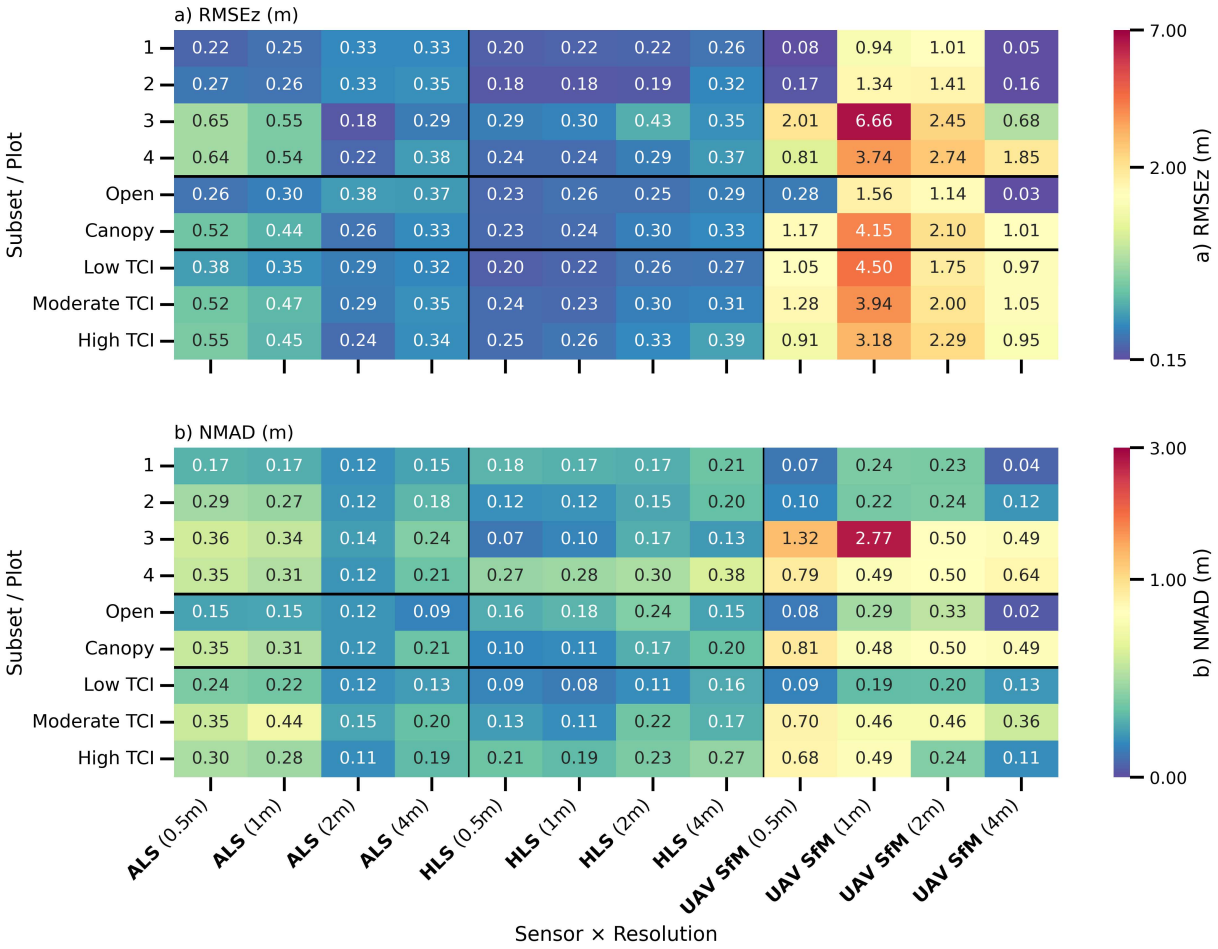


Figure 5.5: Vertical accuracy metrics of DTMs derived from ALS, HLS, and UAV photogrammetry (SfM) for four study plots, two canopy conditions (open and canopy), and three terrain complexity index (TCI) classes (low, moderate, and high), evaluated at four raster resolutions (0.5, 1, 2, and 4 m). Errors are computed relative to full-density UAV LiDAR-derived DTMs. Panels show (a) root mean square error of elevation ($RMSE_z$) and (b) normalised median absolute deviation (NMAD). Lower values indicate higher accuracy.

Conversely, UAV SfM exhibited extreme terrain sensitivity, with roughness coefficients 5 to 23 times higher than ALS (+0.20–1.61 m per unit roughness) and strong correlation with canopy cover ($R^2 = 0.51$) dominating over terrain effects ($R^2 = 0.31$). In the densest canopy plot (99% canopy coverage), SfM errors increased to 6.66 m at 1 m resolution, 7.3 times higher than in open plots (0.91 m). Coarser grids (4 m) mitigated SfM errors somewhat (2.45 m) but not enough to make SfM suitable for forest terrain mapping. These patterns reflect fundamentally different acquisition mechanisms. ALS and HLS compensate for terrain complexity through continuous sampling, whereas SfM's sparse, canopy-limited sampling yields high interpolation errors in complex terrain.

Spatial analysis of canopy and ground-return distributions (Moran's I = 0.75–0.90) revealed highly clustered ground-return voids under dense canopies, confirming that canopy structure drives error distribution and creating systematic mapping challenges in densely vegetated plots (Fig. 5.6).

Table 5.8: Terrain predictors of DTM vertical error (absolute error regression): coefficients (per unit increase in terrain metric) and variance explained (R^2) for each sensor (ALS, HLS, SfM) and spatial resolution (0.5, 1, 2, 4 m). Slope coefficients are expressed per 1% increase in slope gradient.

Res.	Sensor	R^2	R[a]	p	S[b]	p	TWI[c]	p	ME (m)	n pixels
0.5m	ALS	0.18	+0.16 m/unit	***	+0.001 m/%	***	+0.03 m/unit	*	0.35 m	14,620
	HLS	0.15	+0.01 m/unit	ns	+0.001 m/%	*	+0.01 m/unit	***	0.19 m	14,620
	SfM	0.19	+0.30 m/unit	**	+0.001 m/%	ns	+0.12 m/unit	*	0.65 m	14,620
1m	ALS	0.18	+0.07 m/unit	***	+0.001 m/%	*	+0.00 m/unit	ns	0.31 m	3,765
	HLS	0.26	+0.01 m/unit	ns	+0.001 m/%	**	+0.01 m/unit	ns	0.19 m	3,765
	SfM	0.31	+0.73 m/unit	***	+0.011 m/%	ns	+0.61 m/unit	***	3.02 m	3,765
2m	ALS	0.22	+0.02 m/unit	***	+0.001 m/%	ns	−0.08 m/unit	***	0.23 m	1,078
	HLS	0.38	+0.02 m/unit	ns	+0.000 m/%	ns	+0.01 m/unit	ns	0.23 m	1,078
	SfM	0.36	+0.20 m/unit	***	+0.010 m/%	***	+0.47 m/unit	*	1.82 m	1,078
4m	ALS	0.23	−0.15 m/unit	*	+0.013 m/%	*	−0.05 m/unit	*	0.26 m	270
	HLS	0.19	+0.04 m/unit	ns	+0.014 m/%	ns	+0.02 m/unit	ns	0.25 m	270
	SfM	0.32	+1.61 m/unit	*	−0.079 m/%	*	−0.05 m/unit	ns	0.56 m	270

[a] Roughness (R): m DTM error per 1-unit increase in roughness.
[b] Slope (S): m DTM error per 1% increase in slope gradient.
[c] Topographic wetness index (TWI): m DTM error per 1-unit increase in TWI.
ME: mean error.
Significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = not significant.

5.4.4 Slope category agreement

Fig. 5.7 illustrates how grid resolution alters stand-level slope statistics and the composition of slope-based categories derived from the full-density UAV LiDAR DTM. From 0.5 m to 30 m, the median slope decreases by about 6–7 percentage points and the 90th percentile by approximately 20 percentage points, indicating that coarser grids smooth out local steepness extremes. Consequently, the steepest category C (≥50%) declines from one-fifth to almost zero, while gentle-to-moderate category A (≤30%) expands from just

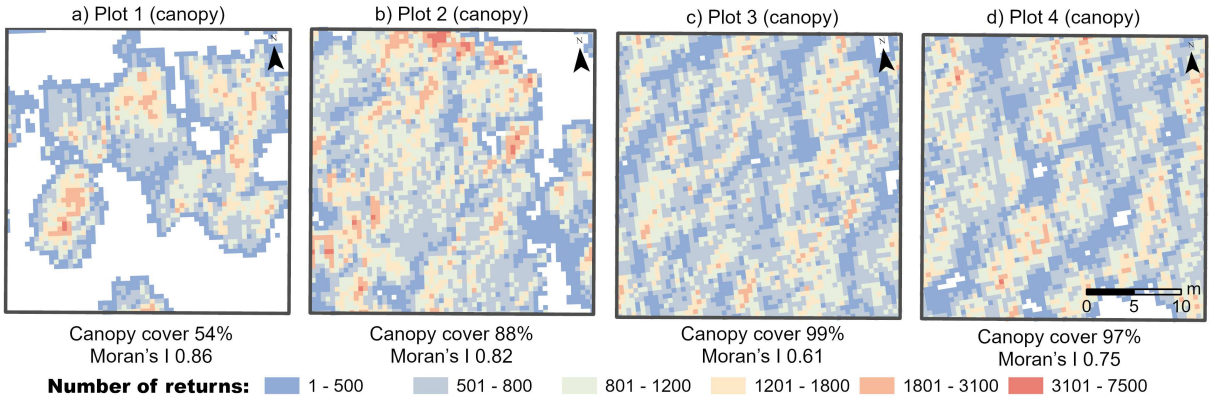


Figure 5.6: Spatial distribution of canopy and ground-return density across the four reference plots derived from UAV LiDAR (0.5 m grid resolution). The panels (a–d) show canopy-return density, expressed as the number of LiDAR returns per 0.5 m cell, together with canopy cover (%) and Moran's I, which quantifies spatial autocorrelation of canopy presence (values near 1 indicate strong spatial clustering, values near zero random distribution).

over half to more than 90% of the area. This indicates that grid resolution alone can substantially alter the mapped extent of slope categories within a stand, independent of the underlying terrain.

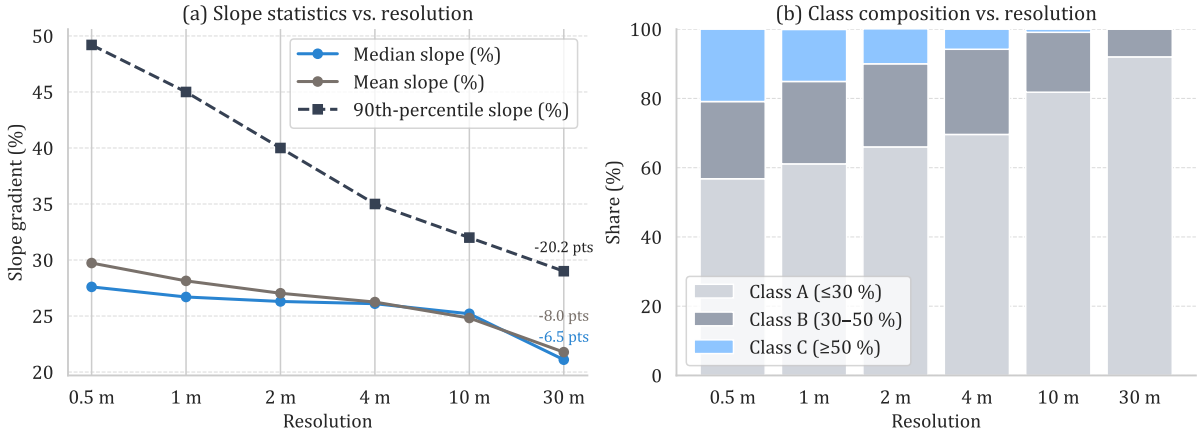


Figure 5.7: Effect of grid resolution on slope statistics (a) and terrain category composition (b) derived from full-density UAV LiDAR. Panel (a) shows median, mean, and 90th-percentile slope gradients across resolutions from 0.5 m to 30 m. Panel (b) displays the proportional area of three slope-based terrain categories: A (≤30%), B (30–50%), and C (≥50%).

To assess how closely the sensor-derived DTMs reproduce the UAV LiDAR–derived slope-based categories, we compared sensor-based categories A, B and C with the UAV LiDAR reference at each grid size (Fig. 5.8). OA increased systematically with coarsening grid size for all sensors (Fig. 5.8a). At 0.5 m, ALS and UAV SfM agreed with the reference for only about 55–65% of pixels, whereas thinned UAV LiDAR DTMs agreed with the reference for roughly 65–75% of pixels. By 2–4 m, OA had increased to approximately 75–95%, with the highest values for LiDAR thinnings at 10–30 pts m^{-2}. At the coarsest resolutions

(10–30 m), LiDAR-based DTMs achieved OA above 90%, whereas the NASA DEM at 30 m reached about 80–85%. However, these high values reflect the absence of very steep category C terrain.

The direction of misclassification showed clear patterns across sensors (Fig. 5.8b,c). At fine grids (0.5–2 m), ALS and especially UAV SfM exhibited relatively high overestimation and underestimation, with both UE and OE often exceeding 10–20%. Thinned LiDAR products showed much smaller directional errors, typically in the single-digit range by 2–4 m. Category-specific accuracies confirmed that category A was consistently mapped most reliably across sensors and resolutions, while category C had the lowest agreement, particularly for ALS and UAV SfM at 0.5–2 m. Accuracy also varied with TCI, with lower OA and higher UE/OE in high-TCI areas compared with low-TCI terrain.

In terms of terrain representation, this indicates that if fine-scale (0.5–2 m) per-pixel A/B/C categories are required, both sensor choice and LiDAR ground-point density become critical. If coarser grids (4–10 m) are used, ALS and relatively sparse LiDAR data agree well with the UAV LiDAR reference. However, this agreement partly reflects the smoothing of many very steep patches visible at finer resolutions. Because these accuracies are obtained from terrain products with very different data and processing requirements, it is also important to compare the data volume associated with each sensor–resolution combination.

5.4.5 Data volume of derived terrain products

Fig. 5.9 summarises the data volume of terrain products for each sensor–resolution combination, expressed as MB per hectare of point clouds, DTMs, and slope rasters. For ALS, data volumes remain below 1 MB ha^{-1} across all grid sizes, and the NASA DEM at 30 m is similarly lightweight (about 1.1 MB ha^{-1}). Thinned UAV LiDAR products span a wider range, from approximately 0.2–0.3 MB ha^{-1} (1 pt m^{-2} at 2–30 m grids) up to roughly 6–7 MB ha^{-1} (30 pts m^{-2}). UAV SfM is consistently the heaviest dataset among the tested products at around 34 MB ha^{-1}, whereas full-density UAV LiDAR products are by far the largest, approaching 300 MB ha^{-1} irrespective of grid size, because the point cloud dominates total file size.

At the stand scale considered here, LiDAR grids of 2–4 m with ground-point densities of approximately 5–10 pts m^{-2} provided a balanced representation of terrain information relative to data volume. These configurations showed high agreement with the UAV LiDAR slope-based categories (OA $\approx$ 90% or higher, with low UE/OE) while keeping data volumes on the order of 2–3 MB ha^{-1}, substantially lower than UAV SfM and orders of magnitude smaller than full-density UAV LiDAR. ALS and the NASA DEM yield very lightweight products in terms of MB ha^{-1} but represent a smoother version of steep terrain and, at fine grids, show lower agreement and higher misclassification of steep categories.

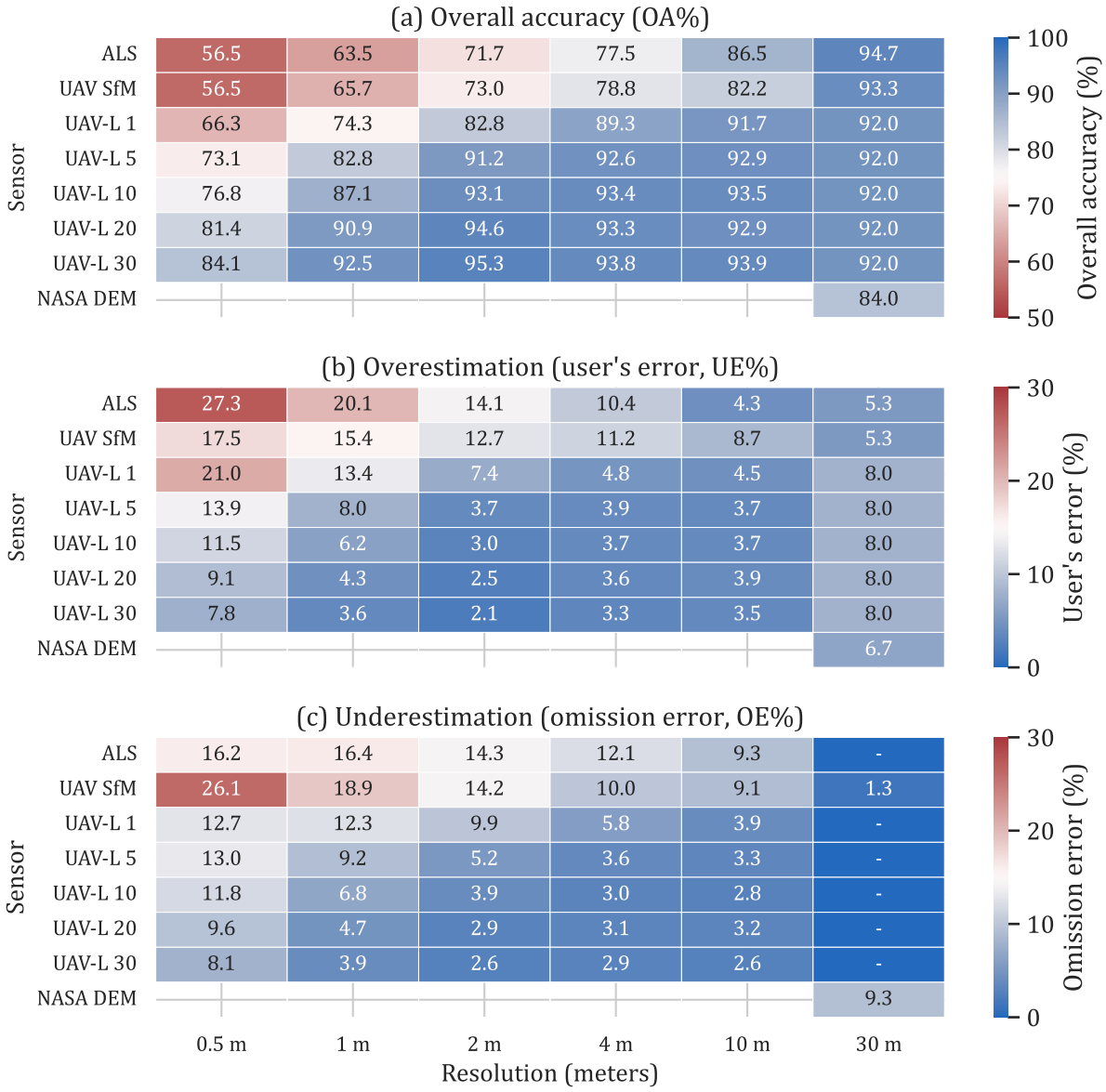


Figure 5.8: Slope-class agreement with the full-density UAV LiDAR reference across sensors and grid resolutions. Heatmaps show (a) overall accuracy (OA, %), (b) overestimation (user's error, UE, %), and (c) underestimation (omission error, OE, %) of aggregated slope-based categories A (≤30%), B (30–50%), and C (≥50%). Rows correspond to sensors (ALS, UAV SfM, and UAV LiDAR thinned to target ground-point densities of 1, 5, 10, 20, and 30 pts m^{-2}, denoted UAV–L 1 to UAV–L 30, and the NASA DEM) and columns to DTM/slope raster resolutions from 0.5 to 30 m.

5.4.6 Slope and TRI across scales

Because slope-based terrain categories (A, B, C) do not differentiate between smooth and highly irregular surfaces, slope was complemented with the TRI. Correlation analysis for the full-density UAV LiDAR DTM showed that raw TRI with small neighbourhoods (R = 2 m) is very strongly associated with slope ($\rho \approx 0.9$) across grid sizes from 0.5 to 4 m (Fig. 5.10a). For larger windows (R = 4–6 m), the correlation with slope remains moderate to high at fine resolutions but declines at coarser grids (10–30 m), where TRI becomes increasingly dominated by broader-scale slope trends. These results indicate that mapping

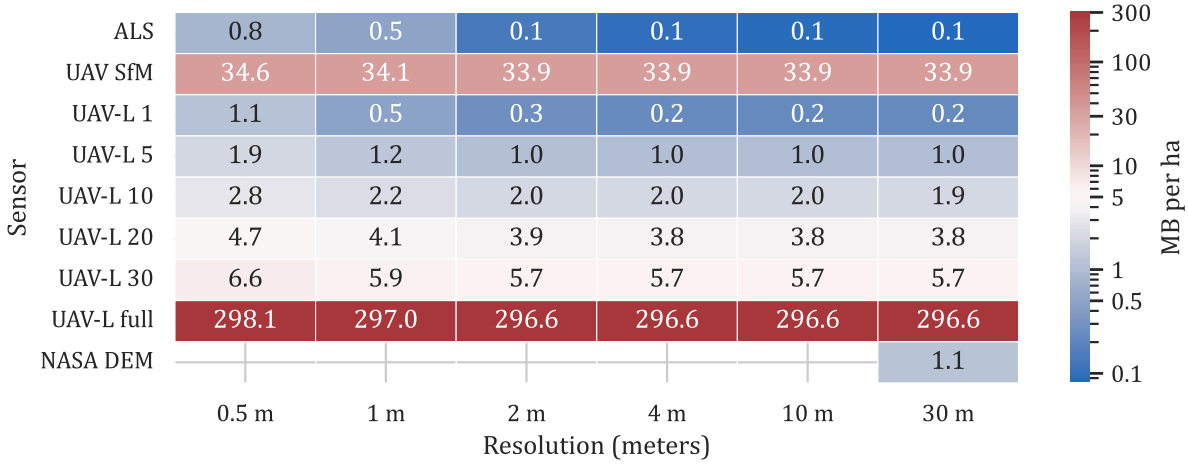


Figure 5.9: Data volume of derived terrain products (MB ha^{-1}) by sensor (rows) and DTM/slope raster resolution (columns). Cells show total file size per hectare of the corresponding point cloud (.las), DTM (.asc, .tif), and slope raster (.tif). UAV LiDAR datasets labelled 1, 5, 10, 20, and 30 correspond to point clouds thinned to 1, 5, 10, 20, and 30 points m^{-2}, respectively. Colours are mapped using a logarithmic ($\log_{10}$) colour scale for visualisation; values shown (labels/units) are the original, untransformed values.

both slope and raw TRI with small neighbourhoods largely reproduces the same terrain information.

To obtain a more slope-independent measure of terrain variability, a detrended TRI was calculated as the residual after removing the linear dependence of TRI on slope. The CV curves (Fig. 5.10b) show that both slope and raw TRI lose variability as resolution coarsens, with the strongest decline between 4 m and 30 m. At 30 m resolution, TRI variability is very low, indicating that contrasts in surface ruggedness are largely smoothed out. In contrast, the detrended TRI retains greater variability. It shows weaker correlation with slope, particularly at finer resolutions and intermediate neighbourhood sizes, where it begins to represent a distinct dimension of terrain variability rather than a simple proxy for slope.

To further examine this relationship, TRI behaviour was analysed within slopes of 30–50% (category B). Within this interval, raw TRI remained strongly coupled with slope (Fig. 5.10a), with higher slopes generally corresponding to higher TRI values and points clustering near the 1:1 line in the raw–detrended comparison. After detrending, this dependence was largely removed, and the distribution of values expanded vertically, covering a wider range of TRI values for the same slope interval (Fig. 5.11a). The correlation between raw and detrended TRI decreased ($R = 0.46$, $R^2 = 0.21$), indicating that detrended TRI captures a component of terrain variability that is not directly explained by slope.

The histograms in Fig. 5.11b further illustrate this contrast. Detrended TRI shows a broader distribution and higher maximum values than raw TRI (up to ~3.0 compared with ~1.3), indicating improved discrimination between smoother and more irregular terrain within category B. Mean values of raw and detrended TRI are similar, but detrending expands the upper tail of the distribution, separating more rugged locations from relatively

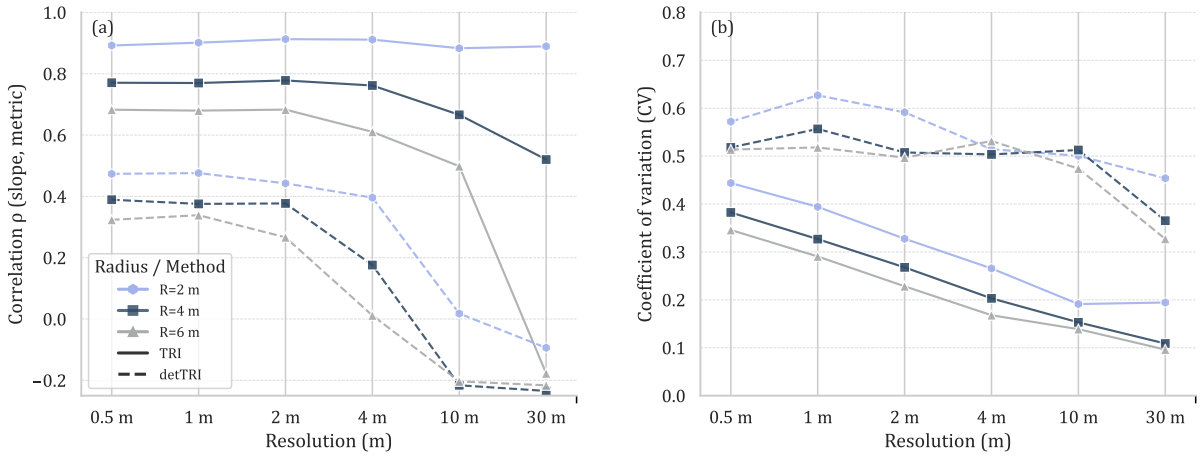


Figure 5.10: Resolution dependence of slope and ruggedness metrics. (a) Pearson correlation between slope and TRI-based metrics across resolutions, including mean slope. (b) Coefficient of variation (CV) of slope and TRI-based metrics across resolutions.

smooth surfaces within the same slope interval. As a result, detrended TRI provides additional information beyond slope and raw TRI alone, allowing steepness and local surface irregularity to be represented as partially independent terrain descriptors.

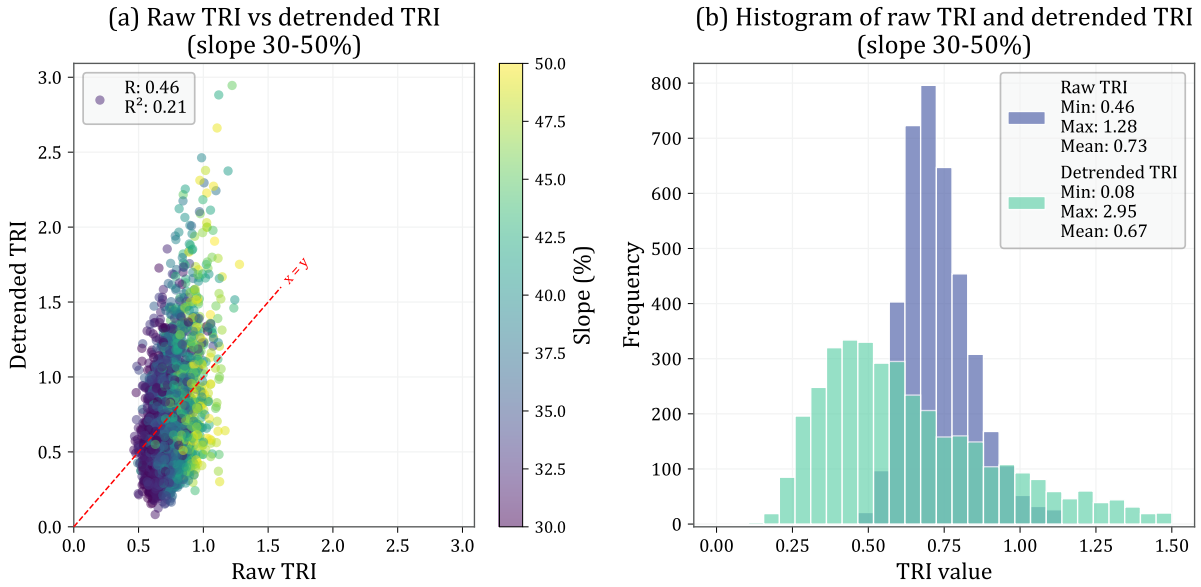


Figure 5.11: Effect of detrending on terrain ruggedness index (TRI) for areas with 30–50% slope (slope-based category B). (a) Relationship between raw TRI and detrended TRI; points are coloured by local slope (%). The dashed red line indicates the 1:1 line ($x = y$). (b) Histograms of raw TRI and detrended TRI, showing the distribution of values before and after detrending. Summary statistics (min, max, mean) are shown in the legend.

5.5 Discussion

This study evaluated how DTMs derived from multiple remote sensing platforms (ALS, UAV LiDAR, HLS, and UAV photogrammetry) influence the reliability and scale dependence

of slope and raster-based terrain-variability metrics in steep, closed-canopy mountain forests. By comparing vertical accuracy, terrain-dependent error patterns, and resolution effects on slope-category agreement, the analysis clarifies how sensor characteristics and grid resolution propagate into terrain indicators commonly used in forest-planning contexts.

The results indicate that HLS and UAV LiDAR provide the most consistent terrain representation under closed-canopy conditions. In contrast, UAV SfM exhibits substantially larger vertical deviations and stronger sensitivity to canopy cover and terrain complexity. Furthermore, LiDAR-derived DTMs gridded at 2–4 m resolution with 5–10 ground points m^{-2} produced stable slope-category outcomes while maintaining manageable data volumes, suggesting a practical balance between terrain representation and implementation feasibility at the stand scale.

In this context, grid resolution and ground-point density determine the extent to which terrain detail is captured and smoothed before slope and terrain metrics are computed. Because stand-level terrain summaries (e.g., dominant slope category or mean slope) are derived mathematically from raster data, their stability depends on how fine-scale variability and extreme values are preserved or smoothed during interpolation and resampling.

Scope of interpretation

It is important to clarify the scope of these findings. This study does not model machine trafficability, predict operational productivity, or quantify obstacle-based roughness in the traditional operational sense [43]. Instead, it evaluates how sensor choice and spatial resolution influence raster-derived terrain indicators that underpin slope-based terrain summaries in forest-planning contexts.

Even when operational decisions rely on aggregated stand-level descriptors, those descriptors are derived from spatial terrain data, and their stability depends on how terrain variability is represented and aggregated. The present results, therefore, address the reliability of the terrain-information pipeline rather than direct operational outcomes.

Sensor performance and error mechanisms

Vertical deviations relative to UAV LiDAR were smallest for HLS, moderate for ALS, and largest for UAV SfM. These patterns are consistent with previous multi-sensor comparisons, which show that active LiDAR generally outperforms passive photogrammetry under forest canopy conditions [210, 164]. SfM errors increased in dense stands, confirming that reliable terrain modelling in closed-canopy forests requires sufficient ground visibility.

Elevation errors were not spatially random but systematically related to terrain and canopy structure. Terrain roughness explained the largest share of elevation deviation, followed by slope and TWI, indicating that irregular micro-relief and limited ground returns

increase interpolation uncertainty. As a result, the largest deviations occurred in steep, rough, closed-canopy areas, conditions where terrain representation is also most challenging. These patterns suggest that DTM reliability varies with terrain conditions and should be interpreted stratified rather than assumed to be uniform across entire stands.

Resolution, scale-matching, and practical feasibility

In forestry terrain analysis, slope, roughness, and ruggedness describe related but scale-dependent properties of a continuous terrain surface. Slope represents the background terrain inclination, whereas roughness and ruggedness describe local elevation variability superimposed on it. Their expression depends strongly on spatial scale: fine grid resolutions preserve small-scale variability, whereas coarser grids smooth it and increasingly absorb it into broader slope trends. In this study, roughness and ruggedness are analysed strictly as geomorphometric indicators derived from gridded DTMs and therefore represent statistical surface variability rather than explicit measurements of individual physical obstacles as defined in traditional object-based roughness classification systems [1].

Grid resolution substantially influenced slope statistics and aggregated terrain summaries. Median slope decreased by 6–7 percentage points between 0.5 m and 30 m resolution, while the 90th percentile declined by approximately 20 percentage points. Consequently, the mapped extent of the steepest category (Category C, $\geq 50\%$) decreased from about 20% at fine resolution to nearly zero at coarse resolution, indicating strong smoothing of local extremes. At the stand scale, grids of 2–4 m combined with 5–10 ground points m^{-2} achieved approximately 90% overall agreement relative to full-density UAV LiDAR, while maintaining limited underestimation of steep terrain. Coarser grids (10–30 m) also showed high agreement, but partly because critical steep segments were smoothed away, whereas very fine grids (0.5–1 m) were more sensitive to local artefacts in slope estimation.

The 30 m NASA DTM was included as a widely available baseline for broad strategic screening. However, the results demonstrate that 30 m products, in general (including resampled UAV LiDAR and ALS DTMs), are not a reliable substitute for stand-level slope mapping in complex terrain, as they systematically suppress localised steep segments relevant to tactical terrain interpretation.

Beyond slope alone, terrain representation can be improved by incorporating surface irregularity metrics. Raw TRI was almost perfectly correlated with slope ($r \approx 0.98$) and therefore contributed little independent information. Detrended TRI reduced this coupling ($r < 0.5$), enabling clearer separation of slope-independent surface variability. Similar slope–ruggedness coupling issues have been documented in geomorphometric research, where residual and slope-independent ruggedness metrics were developed to isolate local terrain variability [79, 84, 211]. These findings support the use of detrended TRI as a slope-independent geomorphometric descriptor of terrain variability in forest environments. Combining slope metrics with slope-independent ruggedness metrics can therefore provide a more complete representation of terrain variability.

Practical implications for forest operations

These findings support a tiered and resource-efficient approach to terrain mapping for forest-planning applications in mountain forests. In practice, terrain conditions in operational planning are often summarised using aggregated stand-level descriptors such as dominant slope categories or mean slope values. However, these summaries are derived from spatially explicit raster terrain data, and their stability therefore depends on how terrain variability is represented in the underlying DTM.

For stand-level terrain characterisation where slope-category delineation is required, LiDAR-derived DTMs provide the most reliable representation of terrain conditions. The moderate-density specification identified here (5–10 points m^{-2} gridded at 2–4 m) captures relevant slope patterns while keeping data volume and processing demands manageable, making it suitable for deployment across larger management areas.

In contrast, ALS and global 30 m terrain products remain useful for broad strategic screening and landscape-level terrain overview but tend to under-represent local steepness in complex mountain terrain. Consequently, they should not be relied upon alone for stand-scale slope-category mapping. Integrating slope maps with slope-independent ruggedness metrics, such as detrended TRI, can further improve terrain description by highlighting surface variability that is not captured by slope categories alone.

Methodological considerations and limitations

This work is based on a single, relatively small study area in a mixed spruce–beech forest in the Italian Alps. The site presents highly challenging operational conditions, with steep slopes, rocky terrain, and a dense, closed canopy, where terrain complexity strongly influences machine mobility, safety, and harvesting costs [212, 209]. While such conditions may amplify modelling errors relative to gentler terrain, they also represent contexts in which accurate DTMs are most consequential for planning and risk assessment. Even so, differences in terrain morphology, soils, and canopy structure in other regions may affect both absolute error levels and relative sensor performance, so the results should be tested in simpler stands and landscape contexts before generalisation.

UAV LiDAR was treated as the reference surface, although it is not error-free. GNSS control under open conditions achieved centimetre-level accuracy, and DJI Terra processing produced low checkpoint errors. However, all other datasets (ALS, HLS, and SfM) were subsequently registered to the UAV LiDAR frame, introducing co-registration uncertainties of approximately 0.07–0.09 m for HLS and 0.37 m for ALS. Consequently, the reported $RMSE_z$ and $NMAD_z$ values represent relative deviations from the UAV LiDAR surface rather than absolute terrain error. Small spatial misalignments may therefore propagate into local slope estimates when raster products are compared. In addition, GNSS positioning under dense canopy degraded to several decimetres for HLS plot corners, highlighting a common limitation of forest GNSS positioning.

Further uncertainty arises from ground classification and surface interpolation. Although consistent filtering parameters were applied, differences in sensor geometry, point density, and canopy structure influence how ground and non-ground returns are separated [213, 214, 215]. Low vegetation, woody debris, or rocks may therefore be classified as ground in one dataset but removed in another, producing local elevation differences that propagate into derived terrain metrics. All high-resolution DTMs were also interpolated using a single mean-based method. While this approach isolates sensor and sampling effects for comparison, mean-based interpolation can oversmooth sharp terrain features and is more sensitive to residual non-ground points and data gaps than adaptive interpolation methods [216, 165, 217, 160].

Additional limitations relate to the derived terrain indicators and their interpretation. The TCI weights were derived from external reference datasets rather than site-specific calibration, and the tercile-based thresholds are therefore best interpreted as comparative descriptors for stratifying terrain variability rather than universal operational breakpoints. Likewise, the slope-based terrain categories (A, B, C) represent simplified terrain descriptors. Actual machine mobility depends on additional factors not represented in the terrain model, including soil bearing capacity, soil moisture, surface cover, and operator practices. Furthermore, the results were not validated against observed machine behaviour such as slip events, winch usage, productivity, or soil disturbance.

The present analysis, therefore, focuses on how sensor choice, point density, and spatial resolution influence the representation and aggregation of raster-derived terrain metrics, rather than directly predicting operational outcomes. Future research should link these terrain descriptors to observations from active forest operations to quantify how DTM resolution and terrain smoothing translate into operational performance and risk.

Key priorities for future work include: (a) extending the analysis to a wider range of forest types and terrain conditions, (b) testing alternative ground-filtering and interpolation methods in complex forests, (c) quantifying acquisition and processing costs for different sensor–resolution combinations, and (d) validating terrain descriptors against machine-based observations of mobility, safety, and soil disturbance.

Broader implications and future directions

Despite these limitations, the results demonstrate that in steep, closed-canopy mountain forests, LiDAR-derived DTMs with moderate ground-point densities and appropriately selected grid resolutions can provide stable terrain descriptors for stand-scale slope categorisation. Combining slope with slope-independent ruggedness metrics offers a more complete representation of terrain variability in forest-planning contexts.

However, these raster-derived geomorphometric indicators complement rather than replace traditional terrain classification systems used in forest operations, which incorporate field-based assessments of obstacles, soil bearing capacity, and surface conditions. While geomorphometric metrics describe spatial variability in terrain form, operational

terrain classifications integrate additional physical and mechanical constraints that influence machine performance.

Future research should extend this work by linking raster-derived terrain descriptors to observations from active forest operations. Integrating terrain metrics with machine-based data, such as routing behaviour, slip events, productivity variation, or soil disturbance, would allow a more direct evaluation of how DTM resolution and sensor characteristics influence operational outcomes.

Further developments could also explore the integration of geomorphometric terrain descriptors with soil- and moisture-related indicators (e.g. depth-to-water indices, soil maps, or seasonal moisture proxies) to better represent the combined factors influencing machine mobility. Finally, multi-site validation across contrasting forest types and terrain morphologies would help assess the transferability of the recommended sensor and resolution specifications under different operational conditions.

5.6 Conclusion

This study evaluated how sensor type, ground-point density, and grid resolution influence DTMs and raster-derived terrain metrics in steep, closed-canopy forest terrain. Using full-density UAV LiDAR as an internal reference at plot and stand scales, HLS provided the most accurate local benchmark, with low and relatively stable vertical deviations across terrain-complexity classes. ALS showed moderate performance, with errors increasing in steeper and rougher terrain and under dense canopy. At the same time, UAV SfM exhibited substantially larger canopy- and terrain-dependent deviations, indicating limited suitability for bare-earth modelling in comparable forest conditions.

At the stand scale, grid resolution strongly influenced slope statistics and slope-category aggregation. Coarsening from 0.5 m to 30 m progressively smoothed steep terrain and nearly eliminated the steepest category (Category C, $\geq 50\%$), altering slope distributions and masking local terrain extremes. LiDAR DTMs gridded at 2–4 m with approximately 5–10 pts m^{-2} achieved high ($\sim 90\%$) agreement with full-density UAV LiDAR-derived slope categories while substantially reducing data and processing demands.

Raw TRI at fine resolution was strongly coupled with slope, whereas detrended TRI captured additional slope-independent surface variability. Combining 2–4 m LiDAR-derived slope categories with detrended TRI therefore provides a consistent geomorphometric framework for representing both steepness and local surface variability within forest terrain.

Overall, the results demonstrate that terrain indicators used in forest-planning contexts are sensitive to sensor characteristics and spatial resolution. Even when planning relies on aggregated stand-level terrain summaries, those summaries depend on how terrain variability is represented and processed at the raster level. This study therefore provides methodological guidance for selecting sensor–resolution combinations that yield stable,

comparable terrain descriptors in complex mountain forests.

5.7 Data availability

The corresponding paper is currently in draft form and will be submitted to a scientific journal in April 2026.

Chapter 6

Data feasibility and operational considerations

Beyond technical accuracy, the use of remote sensing in forest operations also depends on practical factors such as hardware needs, processing complexity, data volume, costs, and time efficiency [218, 219, 220]. These aspects are often treated separately rather than synthesised into a single framework in the remote sensing and forestry literature, yet they are critical for practitioners when choosing and implementing sensor systems. This chapter briefly examines the operational feasibility of different sensor platforms and deployment scenarios in forest operations. It considers survey efficiency, processing and computational requirements, data storage and management, cost structures, and logistical and other operational constraints, to support more informed decisions about which remote sensing technologies to use.

In this chapter, feasibility refers to the practical ability to acquire, process, manage, and apply remote sensing-derived terrain data under real-world forestry constraints. It is treated separately from analytical accuracy and evaluated using indicators of survey efficiency, processing demand, data volume and storage, cost structure, and operational constraints, such as logistics and weather.

6.1 Survey efficiency and data collection effort

Survey efficiency and data-collection effort, which determine how quickly data can be collected across an area, fundamentally shape project feasibility, timelines, and overall cost-effectiveness [221, 222]. Different sensor systems generate data at vastly different rates and require distinct field-deployment strategies. Therefore, understanding these differences is essential for practitioners selecting appropriate technologies for specific operational contexts.

To explore these feasibility trends within the scope of this work, an additional set of data-feasibility metrics was compiled for one of the study sites described earlier, where four remote sensing technologies were applied (see Chapter 5). The comparison of these four sensor approaches clearly illustrates the efficiency differences (Table 6.3). On the same forest site (~5.44 ha), UAV LiDAR and photogrammetry each required ~39 minutes

of field acquisition, whereas handheld LiDAR (HLS) needed 41 minutes to survey only four small plots (0.36 ha total). In contrast, the ALS data used here were acquired as part of a broader airborne campaign covering a ~2×2 km tile (~400 ha) and were subsequently clipped to 5.44 ha for analysis. Therefore, the ALS acquisition time is not directly comparable to that of on-demand UAV or terrestrial surveys.

At operational scales of 100–200 ha, ALS typically offers substantially higher area coverage per unit time than UAV-based LiDAR or photogrammetry, and substantially higher efficiency than terrestrial methods. This speed advantage, however, comes at the cost of lower ground-point density (and often reduced under-canopy detail) compared with UAV and handheld systems, which can provide denser local measurements but require markedly greater field time and processing effort to cover the same area.

Plot-level studies of terrestrial methods show similar patterns. Piermattei et al. [223] compared low-cost terrestrial SfM photogrammetry, TLS, and conventional manual inventory on the same forest plots and found that, while SfM can match TLS and manual measurements in field time under favourable conditions, its efficiency drops sharply in complex terrain. Its processing times can extend to many hours per plot, making routine operational use challenging. Similarly, Pavelka et al. [220] concluded that conventional terrestrial approaches, such as TLS, are 'practically unusable' for extensive operational forestry inventories due to cumulative effort and logistical hurdles. At the same time, mobile and unmanned systems were presented as "economically and operationally attractive" alternatives due to their rapid data acquisition and manageable processing demands.

6.2 Processing demands and computational requirements

The computational demands of different remote sensing systems present another significant constraint for remote sensing projects. Because hardware and computational requirements often scale more than linearly with dataset size and output resolution, they become critical factors in infrastructure investment and project scheduling [**?**]. Processing frequently represents the operational bottleneck, the slowest step that determines the overall project timeline. Understanding processing requirements early in project planning is therefore essential.

This challenge is compounded by the rapid growth of remote sensing data over the past few decades. A global example is NASA's Earth Observing System Data and Information System (EOSDIS), which currently adds on the order of $>$ 100 terabytes (TB) of Earth observation data per day to its archives. Industry and agency summaries suggest that, when commercial operators and multiple space agencies are included, global remote sensing systems collectively generate data volumes on the order of several hundred terabytes per day [224, 225, 226]. This expansion is driven by improvements in sensor resolution and revisit frequency, as well as by the expansion of open-access data-sharing initiatives. At

the scale of forestry and agriculture, UAV- and LiDAR-based remote sensing have likewise moved from niche applications to mainstream tools over the past decade, with a rapid increase in published studies and widespread use for forest inventory, biomass estimation, and canopy mapping [218, 227]. Such volumes and complexities place remote sensing firmly within the domain of *Big Data*, where data-intensive workloads often exceed the capacities of conventional desktop computing environments [224, 228].

While these figures refer to large-scale archives, similar computational constraints were evident even in the relatively small, routine field surveys considered in this thesis. At our study site (Chapter 5), peak hardware usage and processing time were recorded for three operational workflows: UAV LiDAR data processed in DJI Terra [199], UAV RGB imagery processed in Agisoft Metashape Professional [122], and HLS scans acquired with the STONEX X120GO and processed in the vendor software GO Post [200]. All datasets were processed on the same high-end workstation equipped with an Intel Xeon W-2295 CPU (18 cores/36 threads), an NVIDIA RTX 4000 Ada GPU, and 128 GB of RAM, with specifications well beyond typical entry-level systems recommended for remote sensing workloads [122, 200]. Nevertheless, all three workflows saturated CPU capacity (100% utilisation), with GPU load peaking at up to 100% and RAM usage reaching 83% (Table 6.1).

Table 6.1: Peak hardware usage recorded during dataset processing for a mountain forest study site near Cesuna on the Asiago Plateau (Veneto, northeastern Italy). Values are expressed as percentages of system capacity on a workstation equipped with an Intel Xeon W-2295 CPU (18 cores/36 threads), an NVIDIA RTX 4000 Ada GPU, and 128 GB of RAM.

Software	**CPU peak (%)**	**RAM peak (%)**	**GPU load peak (%)**	**GPU memory peak (%)**	**Processing time** (see Table 6.3)
DJI Terra	100	58	100	94	35 min [6.4 min/ha]
Agisoft Metashape Professional	100	89	26	78	6 h (360 min) [66.2 min/ha]
GO Post (STONEX)	100	54	5	81	~5 h 40 min (340 min) (~85 min/plot) [944 min/ha][1]

Photogrammetric processing proved considerably more RAM-intensive than LiDAR workflows in our case study. The SfM workflow in Agisoft Metashape Professional reached a peak RAM usage of 89 % of system memory (approximately 114 GB), compared with 58 % (around 74 GB) for UAV LiDAR processing in DJI Terra and 54 % (around 69 GB) for HLS processing in GO Post (Table 6.1). Sustained CPU utilisation during the main reconstruction steps was also markedly higher for the SfM workflow than for handheld LiDAR processing (approximately 70 % vs 25 %), highlighting that dense image matching and point cloud generation can drive prolonged high-load periods on the workstation. These patterns are consistent with other studies, where high data volumes from under-canopy

UAV photogrammetry, and high-density TLS or UAV LiDAR point clouds lead to long processing times and intensive data management, which in turn limit how many plots or flights can be handled operationally in a given campaign [223, 229, 230].

Similarly, processing times varied widely between workflows, ranging from about 35 minutes for the UAV LiDAR dataset processed in DJI Terra to approximately 6 hours (around 85 minutes per plot) for HLS with manual registration and refinement (Table 6.1, 6.3). Although these absolute runtimes are specific to the high-end workstation configuration used in this study and will vary with hardware, software versions, and parameter choices, they indicate the magnitude of differences that can arise between workflows in practice. This dramatic spread underscores that hardware configuration, software selection, and data preprocessing can significantly affect processing timelines and overall project feasibility.

6.3 Data volume and storage considerations

Beyond acquisition and processing, the volume of data generated and stored represents another critical operational constraint. Data volumes generally increase with point density and output richness, directly affecting field data handling, storage infrastructure, and downstream processing capacity.

The rate at which raw data is generated during field surveys has immediate practical consequences for field operations. In the Cesuna study (Chapter 5), raw data-generation rates ranged from 0.01 GB min^{-1} (ALS) to 0.42 GB min^{-1} (HLS), representing roughly an order-of-magnitude difference between airborne and handheld acquisition. In practice, this means ALS-style datasets are typically straightforward to store, duplicate, and transfer in the field. In contrast, handheld workflows can generate large volumes rapidly, making high-capacity storage media (e.g., external SSDs) and disciplined backup routines essential.

For handheld surveys in particular, data protection becomes part of survey design. Because acquisition may be time-intensive and plots may be difficult to revisit, immediate duplication (e.g., two independent copies on-site) or a consistent end-of-day transfer protocol can be as operationally critical as the scanning itself.

The transition from raw to processed data introduces another layer of complexity. In the Cesuna study, raw data volumes ranged from 0.34 GB for ALS (clipped to the study area) to 17.3 GB for HLS, while processed outputs ranged from 0.04 GB for ALS to 27.3 GB for UAV SfM (complete Agisoft Metashape Professional project; Table 6.3). To compare workflows more systematically, two indicators were derived for each sensor: (a) the raw data rate (RDR), defined as total raw data volume (GB) divided by total survey time (min), yielding GB min^{-1}, and (b) a compression ratio (CR), defined as output data size divided by raw data size (unitless).

In our case, RDR values ranged from 0.01 GB min^{-1} (ALS) to 0.42 GB min^{-1} (HLS), reflecting the much higher data-generation rate of handheld surveys. Compression ratios

also varied strongly between methods. ALS showed the lowest CR (~0.13), indicating a substantial reduction from raw data to final products, whereas HLS showed the highest CR (~0.68), meaning that most of the raw volume carried through into processed outputs (Table 6.3). UAV LiDAR and UAV SfM fell between these extremes. Together, RDR and CR highlight that sensor choice directly translates into storage requirements and processing workload, even within the same forest site.

Overall, robust field-side data protocols (redundant storage, routine integrity checks, and secure transfer to office infrastructure) are most critical for high-volume methods, where data loss would be costly or impossible to recover.

6.4 Cost structure and economic feasibility

Sensor choice ultimately depends on weighing costs and benefits, including equipment purchase, processing infrastructure, field staff, and per-hectare project costs [220, 231]. The economics of remote sensing technologies have shifted dramatically over the past decade. Ten years ago, a widely used drone such as the DJI Phantom 3 Professional retailed for around €1,399 in Europe (2015), dropping to €1,199 by 2016 [232]. Today, entry-level drones such as the DJI Mini series are available for approximately €299. At the same time, advanced consumer models range from €799 to €1,129, making these platforms broadly accessible [232, 233]. Although advanced UAV platforms and LiDAR payloads remain costly, drone-based LiDAR systems typically represent a substantial investment (often in the five-figure range), depending on sensor specifications and whether integrated GNSS/IMU components and software are included. Nevertheless, compared with earlier generations of crewed airborne LiDAR surveying, these systems are now more accessible for stand-scale operational deployments.

Despite declining hardware prices, project budgets are increasingly dominated by soft costs, such as personnel time, data processing, and analysis [220, 234, 235]. Inflation and rising wages in the geospatial sector have increased labour costs in recent years [236, 237], making workforce and training costs among the most important financial considerations for remote sensing operations at all scales.

One such training, crucial for forestry and other remote sensing research projects, is obtaining the necessary drone pilot certification, which requires candidates to complete comprehensive theoretical and practical training, often at significant expense. In the EU, basic drone licences generally cost between €96 and €330, while advanced operator training can range from €1,000 to over €2,450 depending on the level and country [238]. The process is time-intensive and often includes both theoretical and practical exams, with proficiency requirements in some cases on par with those for pilots of manned aircraft. For example, in South Africa, intensive drone pilot training overlaps with modules from commercial aircraft courses, with total costs reaching around €2,000 [239, 240].

When considering the technology itself, the principle is usually quite simple: the choice

of sensor or platform must align with the project's scale and objectives [220, 221, 241]. Notably, the cost per hectare, which reflects both equipment and operational expenses, varies significantly with the total area surveyed [231, 241].

Cost-efficiency changes non-linearly with survey extent because both airborne and UAV surveys include fixed "entry" costs (e.g., planning, mobilisation, permits, and processing setup) that are amortised over larger areas [242]. A practical illustration is provided by the *For.Mo.Vi.* cost assessment, which reports indicative market-based cost comparisons (derived from interviews with practitioners and intended as order-of-magnitude guidance rather than absolute price references) for multiple platforms across 1–1000 ha scenarios. For the specific comparison of airborne LiDAR versus UAV multi-rotor in an "ideal forest area" scenario, the report notes that UAV multi-rotor is more economical up to ~200 ha, costs are broadly comparable between ~200 and ~550 ha, and above ~550 ha airborne LiDAR becomes on average more economical per hectare (Table 6.2). This supports the general principle that platform choice is scale-dependent, while the exact breakpoints depend on site logistics, survey specifications, and local cost structure.

Table 6.2: Comparative cost efficiency of UAV multi-rotor versus airborne LiDAR surveys across different forest area extents, based on operational assessments in the For.Mo.Vi. project [242].

Survey extent (ha)	Cost efficiency comparison	Preferred platform
< 200 ha	UAV multi-rotor has lower setup costs and is typically more economical per hectare	UAV Multi-rotor
200–550 ha	Costs are broadly comparable; choice depends on specific site logistics and data needs	Case-dependent
> 550 ha	Airborne LiDAR fixed costs are amortised; becomes more economical per hectare	Airborne LiDAR

At our study site (Chapter 5), both total cost and surveyed area showed considerable variation across sensing technologies. Table 6.3 presents key feasibility parameters from four remote sensing methods, illustrating dramatic differences in per-hectare costs, data size, processing time, and software requirements. In this case, HLS was more than 60 times as expensive per hectare as ALS and over 13 times as expensive as UAV LiDAR. UAV photogrammetry and LiDAR delivered intermediate costs, remaining several times above the traditional ALS approach but still far less expensive than handheld solutions. Additionally, these advanced platforms generated far larger and denser data outputs, enabling richer analysis but also introducing new challenges for data processing and storage.

In contrast to the market-based contracting scenarios summarised in [242], the per-hectare values in our case study should be interpreted as indicative, because ALS data were

acquired as part of a larger-area campaign and clipped to the study site, whereas UAV and handheld datasets reflect on-demand deployments.

Table 6.3: Detailed comparison of survey, data size, and processing metrics by method.

Metric / method	**ALS (2022)**	**UAV LiDAR**	**UAV SfM**	**HLS**
Platform	Helicopter LiDAR	DJI Matrice 300 RTK + Zenmuse L2	DJI Matrice 300 RTK + Zenmuse L2 RGB (20 MP, 4/3" CMOS)	STONEX X120GO
Survey area (ha)	~400 ha (one 2×2 km tiles); 5.44 ha clipped for analysis	~5.44	~5.44	4 plots (~0.36 ha)
Survey time	~16 min [0.08 min/ha]	~39 min [7.2 min/ha]	~39 min [7.2 min/ha]	~41 min (avg. 9.5 min/plot) [all: 114 min/ha][1]
Raw data size (D_{raw})	0.34 GB (full) [0.00084 GB ha^{-1}]	8.9 GB [1.63 GB ha^{-1}]	5.4 GB (images) [0.99 GB ha^{-1}]	17.3 GB (all), ~3.9 GB (plot) [all: 48.1 GB ha^{-1}][1]
Raw data rate (RDR)[4]	0.01 GB min^{-1}	0.23 GB min^{-1}	0.14 GB min^{-1}	0.42 GB min^{-1}
Processing time (T_{proc})	~120 min [0.6 min/ha]	35 min [6.4 min/ha]	6 h (360 min)[2] [66.2 min/ha]	~5 h 40 min (340 min) (~85 min/plot) [944 min/ha][1]
Output data size	0.036 GB (clipped) [0.007 GB ha^{-1}]	~48.5 GB (all) [8.92 GB ha^{-1}]	~27.3 GB (full Agisoft project[2]) [5.02 GB ha^{-1}]	~25.5 GB total (~6.4 GB/plot[3]) [70.8 GB ha^{-1}]
Used for analysis (point cloud size)	0.036 GB [0.007 GB ha^{-1}]	7.7 GB [1.42 GB ha^{-1}]	19.3 GB (HQ) [3.55 GB ha^{-1}]	~15.65 GB total (~3.91 GB/plot) [43.5 GB/ha (based on ~0.36 ha)][1]
Processing efficiency index (PEI)[5]	0.0014 GB min^{-1} [0.084 GB h^{-1}]	0.254 GB min^{-1} [15.2 GB h^{-1}]	0.015 GB min^{-1} [0.90 GB h^{-1}]	0.051 GB min^{-1} [3.06 GB h^{-1}]
Compression ratio (CR)[6]	0.128	0.184	0.198	0.680

(continued on next page)

Table 6.3 continued from previous page

Metric / method	ALS (2022)	UAV LiDAR	UAV SfM	HLS
Software used	—	DJI Terra	Agisoft Metashape Professional	GOapp + GOpost
License	—	2.070	3.184	Included
Sensor price (€)	n/a (not disclosed; quote required)	≈ 27,030[7]	≈ 13,500[8]	≈ 26,000[9]
Total cost (€)	n/a	≈ 29,100 [5,352€/ha][7]	≈ 16,684 [3,317€/ha][8]	≈ 26,000 [72,222€/ha][9]

[1] For handheld LiDAR, four plots cover 0.36 ha; therefore, per-ha values are based on small plot extrapolation.
[2] Includes aligned images, dense cloud, DTM and GCP references.
[3] Includes export as raw point cloud (.las), point cloud with dense optimisation (.las), and odometry data (.txt).
[4] Raw data rate (RDR) = Total raw data (GB) / Total survey time (min).
[5] Processing efficiency index = Total raw data size (MB) / Processing time (min).
[6] Compression ratio = Output data size (MB) / Raw data size (MB).
[7] DJI Matrice 300 RTK base platform (source: https://store.dji.com/), DJI Zenmuse L2 LiDAR payload (source: https://store.dji.com/), DJI BS65 charging station (source: https://gpsglobalsolutions.com/).
[8] Source: https://gpsglobalsolutions.com, https://www.agisoft.com/buy/licensing-options/.
[9] Source: https://shop.motmould.com/en/303-stonex-x120go-slam-laser-scanner.html.

Most UAV LiDAR, photogrammetry, and handheld systems rely on specialised, vendor-specific software, such as DJI Terra for UAV LiDAR and Agisoft Metashape Professional for photogrammetry, to process their proprietary raw data (Table 6.3). Licensing these programs typically costs between €2,000 and €3,000, with pricing models ranging from annual subscriptions to one-time purchases. The need for dedicated software not only incurs significant expense and operational complexity but also limits data compatibility and cross-platform data exchange. As such, technology selection has significant implications for project budgets, data richness, and workflow flexibility in forestry mapping. Understanding this complete cost structure, beyond just sensor hardware, is critical for practitioners making decisions about remote sensing technology.

6.5 Logistics and operational constraints

Beyond sensor acquisition and personnel costs, logistics play a crucial role in the feasibility and environmental footprint of remote sensing operations [243, 244]. Transporting large sensors and battery systems, such as the DJI Matrice 300 RTK and Zenmuse L2 LiDAR, requires specialised cases and often multiple vehicles for field campaigns. For example, the battery station for the Matrice 300 RTK can weigh up to 8.4 kg when empty and over 15 kg when fully loaded, resembling a thick suitcase and substantially increasing transportation energy use and emissions [245], as well as the personnel transported to field sites. Recent EU reports highlight that transport logistics, including vehicle use and fuel consumption,

can significantly contribute to the carbon footprint of field operations [243, 244]. However, the full environmental impact, including field transportation emissions, increased fuel use, and the energy required for charging large battery systems, remains poorly studied in the current remote sensing literature. As these logistical demands grow with larger platforms, there is a clear need for further research into their sustainability and environmental footprint.

6.6 WEATHER AND SEASONAL CONSTRAINTS

Different sensors have distinct weather tolerances that affect project planning and timelines. ALS operations require cloud cover <10% and wind speeds <12 m/s [246]. UAV systems typically cannot operate above 10 m/s wind or during precipitation [247]. HLS is not directly weather-dependent, but becomes logistically constrained during wet conditions. Terrain accessibility, wet vegetation, and muddy ground significantly reduce survey efficiency. This means that, for time-critical projects, weather risk can also influence technology choice.

6.7 AN EXPLORATORY FEASIBILITY FRAMEWORK

To illustrate feasibility trade-offs in a transparent and reproducible manner (rather than proposing a generalised cost model), we computed three exploratory feasibility indices from Table 6.3: an operational feasibility index, a hardware-cost feasibility index, and an overall feasibility index. Equal weights are used as a transparent baseline. Alternative weightings would be appropriate under different stakeholder priorities.

Operational feasibility was quantified from three lower-is-better metrics expressed per hectare: survey time (T_{survey}), processing time (T_{proc}), and processed output data volume (D_{out}). Each metric was transformed using reverse min–max normalisation,

$$S(x_i) = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}},$$

where $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of metric x across the sensor workflows compared in Table 6.3. Because min–max normalisation is sensitive to the range of values in the comparison set, the resulting indices should be interpreted as relative within this case study rather than as transferable absolute scores. The operational feasibility index was then computed as the arithmetic mean:

$$F_{\text{op},i} = \frac{1}{3}\left(S(T_{\text{survey},i}) + S(T_{\text{proc},i}) + S(D_{\text{out},i})\right).$$

Hardware-cost feasibility was computed analogously from hardware cost (C_{hw}):

$$F_{\text{cost},i} = S(C_{\text{hw},i}),$$

and overall feasibility was defined as the average of operational and cost feasibility:

$$F_{\text{overall},i} = \frac{1}{2}\left(F_{\text{op},i} + F_{\text{cost},i}\right).$$

Accuracy was assessed using the median slope concordance correlation coefficient (CCC) at 1 m resolution relative to a high-density UAV LiDAR reference dataset. The reference was assigned a score of 1.0, and other sensors were scaled accordingly:

$$A_r = 1, \qquad A_i = \frac{\text{CCC}_i}{\text{CCC}_r}.$$

Because ALS hardware/acquisition costs were not available as a citable figure for the third-party campaign data used here, ALS is shown only for accuracy and operational metrics and is excluded from the hardware-cost and overall feasibility indices (Fig. 6.1).

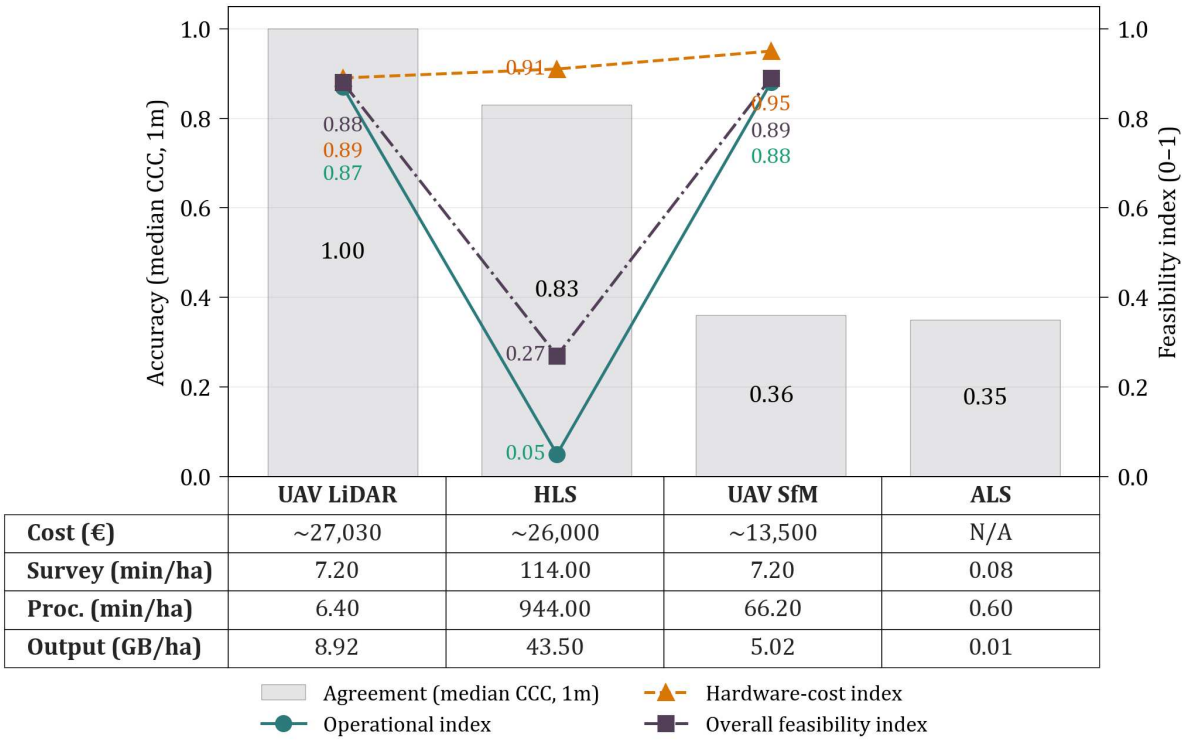
6.8 Synthesis and key message

Data feasibility encompasses far more than sensor accuracy. This chapter examined five interdependent operational factors: survey efficiency, processing demands, data volume, costs, and logistics, to understand why technology selection cannot be reduced to a single metric. No universally optimal platform exists, and research on these operational dimensions remains limited. This gap highlights the need for practical feasibility frameworks that support technology selection in forest management.

Our case study demonstrates this principle concretely. Processing time, personnel demands, logistical complexity, and weather robustness often prove equally or more constraining than nominal equipment cost. A technology may appear inexpensive but become infeasible if it generates data that requires extensive processing, specialised software, or operational effort beyond available resources. Fig. 6.1 shows that no platform ranks first across all dimensions. UAV SfM photogrammetry performs well operationally but is less robust under dense canopy where image matching and ground visibility degrade. HLS achieves high local accuracy but is constrained by slow survey rates and extensive processing demands. ALS is efficient for large-area planning, but lower ground-point density reduces under-canopy accuracy relative to denser systems.

Sensor selection should therefore be tailored to the operational context. For large-area strategic planning ($>$ 200 ha), ALS offers cost-effectiveness and rapid coverage. At stand-to management-unit scales (10–100 ha), UAV systems balance speed, accuracy, and cost. For detailed plot-level assessment ($<$ 10 ha), handheld LiDAR provides the highest accuracy but requires extended field and processing time. The underlying trade-off is inescapable: practitioners must balance area coverage, time-to-result, required accuracy, processing capacity, budget, and operational constraints, including weather and logistics.

Future research should develop practical decision tools that integrate survey time, processing requirements, storage needs, and total project costs across diverse forest types



	UAV LiDAR	HLS	UAV SfM	ALS
Cost (€)	~27,030	~26,000	~13,500	N/A
Survey (min/ha)	7.20	114.00	7.20	0.08
Proc. (min/ha)	6.40	944.00	66.20	0.60
Output (GB/ha)	8.92	43.50	5.02	0.01

Figure 6.1: Accuracy and proof-of-concept feasibility trade-offs across sensor platforms (Cesuna). Bars show slope accuracy (median CCC at 1 m; relative to UAV LiDAR). Lines show feasibility indices (0–1): operational feasibility combines per-ha survey time, processing time, and output data volume (reverse min–max); hardware-cost feasibility reflects relative hardware affordability where citable prices were available; overall feasibility is the mean of operational and cost indices. ALS cost was not citable for the third-party campaign dataset, so ALS is excluded from the cost and overall indices (N/A) but retained for accuracy and operational metrics. The table reports the per-ha inputs.

and regulatory contexts. Such frameworks would enable forest managers to estimate feasibility and costs prior to platform selection, thereby reducing procurement errors and improving technology adoption.

Chapter 7

Discussion

The primary aim of this research was to evaluate the potential of remote sensing–based terrain analysis for objectively quantifying terrain attributes relevant to mechanised harvesting planning, and to assess their reliability and operational feasibility across contrasting forest environments, sensor types, and spatial scales.

To provide a coherent conceptual foundation, the dissertation adopts the three core components of terrain classification systems (TCS) as an organising framework: soil conditions, surface roughness, and slope [42, 43]. Each pillar was addressed through a distinct empirical study: stump influence and root reinforcement as proxies for ground support and deformation response (Chapter 2), obstacle detection for roughness estimation (Chapters 3, 4), and the influence of sensor choice and spatial resolution on DTM-derived slope and terrain-complexity metrics (Chapter 5).

Collectively, these studies demonstrate that remote sensing enables quantitative evaluation of machine-induced surface deformation (e.g., rut depth), surface roughness, steepness, and terrain complexity, and can support proxy-based assessment of ground-condition variability. However, the accuracy and interpretability of derived metrics remain context-dependent, varying with sensor characteristics, spatial resolution, processing choices, and underlying terrain complexity and vegetation conditions.

The remainder of this chapter relates the findings to the three terrain pillars (Section 7.1), synthesises them into a planning-oriented framework (Section 7.1.4), and discusses methodological, operational, and feasibility implications (Section 7.2). It concludes by clarifying the scope of validation (Section 7.3), outlining limitations and uncertainty (Section 7.4), and proposing future research directions (Section 7.5).

7.1 Quantifying the fundamental terrain attributes

7.1.1 Soil conditions: modelling the impact of root reinforcement

In several operational terrain-classification and planning frameworks, ground conditions are represented as polygon-level attributes, including soil type and moisture-related indicators [1]. However, this approach can overlook spatial variability at fine scales, particularly in soft soils, where localised reinforcement from root systems can significantly

alter the soil response to machine loading. Direct measurement of soil strength remains challenging using remote sensing alone because the controlling properties are largely subsurface.

Chapter 2 investigated whether tree stumps and their associated root networks could serve as quantifiable proxies for relative ground-support capacity. By integrating UAV-derived DTMs and stump locations into a spatial decay model, the study achieved R^2 values of 0.69–0.85 for observed rut depths across different machine passes and trail configurations. The key finding was that higher cumulative stump influence was associated with lower rut depth (i.e., reduced deformation), indicating that stump proximity is a relevant explanatory factor alongside soil moisture and texture when evaluating site trafficability for machine operations.

In operational planning, terrain-derived wetness indicators (e.g., DTW, TWI) are widely used in the literature and in operational mapping to identify areas with elevated rutting risk [23, 109]. The present results suggest that adding a stump/root proxy can improve within-stand differentiation of rutting susceptibility in contexts similar to those of this study by accounting for local reinforcement effects. Practically, this supports a more detailed within-stand delineation of zones with lower rutting susceptibility and could be incorporated into routing decisions in peatlands when combined with wetness information.

However, several limitations constrain the current formulation and its transferability. First, the root influence variables (CI, RRV) were modelled as exhibiting a simplified, symmetric decay with distance from the stump [114]. In reality, root systems are typically anisotropic or direction-dependent, with architecture shaped by species, stand age, soil properties, and site-specific environmental gradients [115]. On sloped terrain, root systems often exhibit asymmetric development, with stronger anchorage on the downslope side and preferential lateral growth influenced by slope aspect and prevailing wind direction [248, 249]. This anisotropy means that the actual zone of root reinforcement may be elliptical or directionally biased rather than radially symmetric [114]. Consequently, the simplified circular-buffer assumption used here may underestimate reinforcement in some directions and overestimate it in others.

Although recent TLS-based quantitative structure models have enabled detailed allometric relationships between stump diameter, root system volume, and linear root length for conifer root systems in boreal forests [250], such formulations remain largely species-specific. They are not yet parameterised in terms of distance from the tree or the soil-displacement zone relevant to machine–soil interaction. Moreover, empirical studies of coarse root architecture across Norway spruce, Scots pine, and birch demonstrate significant variation in radial spread, branching intensity, and rooting depth between species and sites [115], underscoring the need for spatially explicit root architecture models that can be coupled with terrain and trafficability assessments.

In addition, although the developed model successfully predicts rut depth as a mechanical outcome, it does not account for the long-term biological consequences of soil defor-

mation. These include changes in soil structure, aeration, and hydrology that can alter root functioning, tree stability, and stand resilience over time. Deep rutting can severely restrict root aeration, increase mechanical impedance, and impair natural drainage [251, 252], although growth responses to such soil deformation vary considerably among tree species and soil textures [253]. Additionally, mechanical damage around stumps may increase their exposure to pathogens, which can reduce stump vitality and resprouting, as shown for aged oak coppice where stump infection risk depends on cut quality [254, 255]. While this study treats stumps solely as mechanical anchors, future research should examine whether traffic also compromises their biological function.

In relation to RO1 and RQ1 (Chapter 1), the results indicate that remote sensing-derived terrain products can (a) capture spatial patterns of surface deformation at extraction-trial scales and (b) support modelling of machine-soil response when combined with operational variables and stump/root reinforcement proxies. In practical terms, this enables relative rutting-susceptibility mapping and within-stand delineation of higher- and lower-risk zones for routing and mitigation. At the same time, the approach remains context dependent. Transfer beyond the study conditions requires recalibration and additional validation across soil types, moisture states and seasons, and machine configurations.

7.1.2 Surface roughness: discrete detection versus continuous metrics

Beyond bearing capacity and soil conditions, machine mobility is strongly constrained by surface roughness in many operational contexts. However, formalising roughness for remote-sensing–based terrain products is not always straightforward for two related reasons: (a) in forestry, roughness is often expressed as obstacle frequency aggregated over plot- or polygon-scale mapping units, which must be reformulated when roughness is represented at pixel resolution; and (b) the meaning of 'roughness' differs between forestry, geomorphometry and other fields [139, 140, 141, 142, 143, 144], with forestry framing roughness as discrete, mobility-limiting obstacles [1, 43] rather than continuous surface texture. Roughness is also formalised in the road-engineering and forest-road literature as road-surface roughness, which is used to characterise ride comfort, vibration exposure, and maintenance needs on unpaved and forest roads [256, 257].

In remote sensing, roughness is often computed as the statistical texture of a DTM [142]. However, this creates a practical disconnect for forestry applications because standard ground-filtering and DTM interpolation are designed to remove or smooth "non-ground" objects, thereby suppressing or attenuating near-ground objects and microtopographic features (e.g., boulders or small terrain depressions) that act as mobility-limiting obstacles.

Chapters 3 and 4 used the post-fire landscape as a 'textbook' setting to implement the TCS obstacle-based definition of roughness, using UAV photogrammetry combined with

image-based instance segmentation to delineate discrete obstacles. This unobstructed visibility enabled an obstacle-detection rate of 95.6% relative to field observations. Implementing the classification at 3 x 3 m resolution then exposed a key scaling limitation: hectare- or polygon-based obstacle-frequency rules cannot be applied directly at the pixel level, where cells typically contain either zero or one obstacle. This required reformulating the original counting logic into a local density representation. This rescaling can introduce inconsistencies relative to the original plot-based rules, and the thresholds and aggregation logic should therefore be refined and tested further in future studies.

Nevertheless, this reformulation still allowed the production of an initial spatial roughness map that represents obstacle-based roughness as a continuous surface. Additionally, these results were compared with conventional roughness metrics derived from DTMs (e.g., TRI), which describe continuous elevation variability rather than discrete obstacles. The comparison indicated that the two approaches can align in capturing broad patterns of uneven terrain. However, they diverge in their meaning and sensitivity because DTM-based metrics are influenced by ground filtering and interpolation and may not preserve individual obstacles. Thus, even when TRI (or related indices) correlates with obstacle-based roughness, it should be interpreted as a different roughness construct than the obstacle-frequency concept used in forestry TCS.

Chapter 5 used continuous DTM-based metrics (including TRI) in closed-canopy forest terrain, where discrete obstacles are challenging to observe reliably from remote sensing. In this setting, TRI is a practical descriptor of surface variability, but it reflects broader terrain form rather than obstacle impediments. Raw TRI was also strongly coupled with slope at the analysed scales, so it often conveyed similar information to slope. This aligns with geomorphometry work cautioning that some “roughness” indices can mainly reflect steepness unless they are interpreted carefully or detrended [84]. Therefore, roughness metrics should be interpreted in terms of the construct they actually represent (obstacles vs continuous undulation) and the mapping scale. Under canopy, DTM metrics are often the most practically available option. Obstacle-based roughness, however, requires workflows that preserve and classify near-ground objects (e.g., point-cloud–native approaches) rather than standard ground-only DTMs.

Referring to RO2 and RQ2 (Chapter 1), the results show that remote sensing can quantify surface roughness in two distinct, construct-specific ways. Where ground visibility is high, UAV photogrammetry combined with instance segmentation can operationalise the obstacle-based roughness concept used in forestry TCS and support spatially explicit roughness mapping at planning-relevant scales. However, translating plot- or polygon-based TCS counting rules into raster products requires explicit aggregation logic, and the resulting class outcomes are therefore scale-dependent. In closed-canopy conditions, discrete obstacle roughness cannot be assumed to be observable from airborne products. Instead, roughness mapping must either shift to DTM-derived surface-variability constructs or use point-cloud–native approaches (e.g., TLS) for near-ground object separation. Conse-

quently, transferability depends primarily on visibility conditions, the roughness construct required by the decision task, and the chosen spatial support.

7.1.3 Steepness and terrain complexity: the influence of sensor and resolution

Slope is a primary terrain constraint for mechanised harvesting because it affects machine stability and influences the feasible range of harvesting systems. In practice, slope information is often obtained from field-based estimates along representative transects or from mapped raster layers, and then summarised into stand- or polygon-level descriptors such as dominant slope class, mean slope, or a min–max range [1]. This type of summarisation can under-represent short, steep segments and local variations that are operationally important for routing and safety [11].

Chapter 5 showed that this under-representation becomes more pronounced as DTM resolution coarsens, because microtopography is progressively smoothed and the mapped extent of the steepest classes is reduced. Across resolutions, degrading the DTM from 0.5 m to 30 m reduced the mapped extent of the steepest slope classes (e.g., $\geq 50\%$), potentially under-representing localised features most relevant for stability and exclusion planning. This highlights a practical advantage of remote sensing for safety-oriented planning: slope can be represented as a spatially explicit constraint layer rather than a single descriptive stand attribute. In planning terms, the key question shifts from "*what is the stand or site slope*?" to "*where are the local steep segments that require avoidance, alternate routing, or a different harvesting system?*" [11].

Chapter 5 also showed that sensor type, LiDAR ground-point density, and grid resolution jointly control how reliably slope classes are represented under closed canopy. For slope-based terrain mapping, intermediate grid sizes (e.g., 2–4 m) combined with intermediate ground-point densities provided a practical compromise between preserving local steep segments and keeping data volumes manageable for operational use.

Preserving this local detail becomes critical when the objective shifts from general planning to detailed environmental mitigation. The loss of micro-topography at coarser resolutions can obscure the specific local slope breaks and channels that organise water flow. If the goal is to prevent gully erosion on steep terrain, planners need to resolve these small flow paths to place water bars and other measures effectively, a level of detail that can be provided by high-resolution digital elevation models (DEMs) and, in some cases, TLS surveys [258, 259]. Therefore, while coarser resolution may suffice for general layout and system allocation, higher-resolution data becomes essential when detailed mitigation planning is required.

A key limitation of this study is that the analysis is based on a single, relatively small, steep, and densely forested study area. This represents a highly complex terrain setting in which slope and microtopography most strongly influence equipment choice, operational safety, and harvesting costs. Additional uncertainty arises from point-cloud processing

steps, especially ground classification and interpolation under canopy, where sensor characteristics and processing workflows can differ in how rocks, low vegetation, and gaps are handled, and these differences propagate into slope and ruggedness layers. Future studies should test transferability in simpler stands and across different forest types, and quantify the robustness of the results to alternative processing settings. These limitations also clarify a practical point for operational use: the "right" sensor, point density, and grid resolution depend on the application. Across all three terrain pillars (Chapters 3, 4, 5), if derived from remote sensing, it is essential to specify which features must be captured, at what accuracy, and at what spatial support, and then choose data specifications accordingly.

In relation to RO3 and RQ3 (Chapter 1), the results show that slope and terrain-complexity layers derived from DTMs are not directly comparable across sensors and scales under closed canopy. Terrain representation and slope-category outcomes depend on canopy penetration (and thus ground-point availability), the spatial support of the grid, and processing choices in ground classification and interpolation. Consequently, slope-based terrain categories are comparable only when the source sensor, point-density regime, grid resolution, and workflow are explicitly specified and when thresholds are applied with a stated spatial support. Finally, the work shows that ruggedness metrics, such as TRI, are strongly coupled to slope at relevant scales, so the interpretation of terrain complexity should explicitly distinguish slope-dependent from slope-independent surface variability (e.g., through detrending) when ruggedness layers are used alongside slope for planning.

To orient the cross-pillar synthesis, Table 7.1 summarises the empirical chapters by terrain pillar, sensor platform, spatial support, and the main derived outputs.

7.1.4 Synthesis across pillars: from terrain layers to planning decisions

Overall, the empirical studies organised around the three terrain pillars demonstrate that remote sensing enables a shift in terrain assessment from stand-level descriptors toward spatially explicit constraint mapping. Rather than characterising a forest stand by a single slope class, roughness category, or soil type, the three pillars (soil conditions, surface roughness, and steepness) can be represented as continuous or categorical raster layers that explicitly show where operational constraints occur. This changes how terrain information can be used in mechanised harvesting planning.

Across the studies, each pillar captures a distinct, complementary mechanism affecting machine mobility. Soil-condition proxies delineate areas with elevated deformation risk and areas where root reinforcement reduces rutting susceptibility (Chapter 2). Roughness can be mapped as obstacle density in open terrain where obstacles are directly detectable using UAV photogrammetry (Chapters 3), and, where airborne visibility is limited, through direct near-ground obstacle detection in TLS point clouds (Chapters 4). Under closed-canopy conditions where discrete obstacles cannot be reliably delineated, rough-

Table 7.1: Overview of data sources, sensors, and analytical tools used across the empirical studies.

Chap.	Pillar	Sensor(s)	Scale	Key outputs	Software / tools	Primary purpose
Ch. 2	Soil conditions	UAV SfM	~cm–dm	DTM, elevation change maps, stump locations, CI, RRV, shear modulus rasters	Agisoft Metashape, Python (RF)	Rut-depth modelling & proxy evaluation
Ch. 3	Surface roughness	UAV SfM	cm–m	DTM, orthomosaic, TRI, VRM, AR, obstacle instances	Agisoft Metashape, Python, SAM DL framework, ArcGIS Pro, QGIS	Obstacle-based roughness mapping
Ch. 4	Surface roughness	TLS	mm–cm	Semantically classified point clouds	CloudCompare; MATLAB; PointNet++	Near-ground obstacle detection and model development
Ch. 5	Slope & complexity	ALS, UAV LiDAR, SfM, HLS, NASA DEM	0.5–30 m	DTMs, slope, roughness, TRI rasters, TCI	LAStools, FUSION/LDV, GDAL, Python, ArcGIS Pro, QGIS	Scale & sensor sensitivity
Ch. 6	Feasibility	All	—	Data volumes, processing times	Logs, scripts	Operational feasibility analysis

ness can also be represented as terrain surface variability derived from DTMs (Chapter 5). High-resolution slope layers identify short, steep segments that are essential for machine stability and for selecting a harvesting system (Chapter 5).

Considered jointly, these layers support a planning perspective that differs from polygon-based terrain classification. Therefore, instead of asking "*what is the dominant terrain/slope category of this stand?*", the relevant question becomes "*where do local combinations of wet ground, roughness, and steepness necessitate avoidance, mitigation, or an alternative harvesting system?*". A broadly trafficable stand may still contain short, steep banks or obstacle clusters that constrain routing. In contrast, some soft-soil areas may remain operable where stump density provides local reinforcement, particularly when considered together with wetness indicators.

In this sense, remote sensing does not replace established terrain concepts, but refines them by explicitly representing within-stand variability. The three pillars function as complementary constraint layers that can be queried at decision-relevant scales: at strategic scales for system allocation and access planning, and at tactical or operational scales for corridor selection, machine routing, and the identification of local exclusion or mitigation requirements.

At the same time, the interpretation of these layers depends on sensor characteristics, grid resolution, and processing choices. Because mapped constraints are scale- and method-dependent, terrain products should be selected to match the planning task. In practice, this requires specifying which constraints must be represented, at what spatial scale, and with what acceptable uncertainty, and then selecting sensors and processing workflows that meet those requirements. Table 7.2 summarises the main practical implications of this synthesis by linking each TCS pillar to suitable sensing approaches, the dominant scale dependencies, and key feasibility constraints.

Table 7.2: Synthesis of the TCS pillars used in this work with respect to suitable sensing approaches, scale dependencies, and feasibility considerations for mechanised harvesting planning.

Terrain pillar	Primary sensing approaches	Key scale dependencies	Feasibility considerations
Soil conditions	UAV photogrammetry	Effective at fine scales for relative rutting risk; requires recalibration across soils, seasons, and machines	Requires field calibration; moderate data volume; limited transferability without site-specific data
Surface roughness	UAV photogrammetry (open terrain); TLS / mobile LiDAR (near-ground obstacles); DTM-derived metrics	Obstacle-based methods sensitive to visibility; DTM metrics capture broader surface variability	High data volume for point-cloud methods; obstacle detection constrained by vegetation and canopy
Slope and terrain complexity	HLS; UAV LiDAR; UAV-SfM (only where ground visibility is sufficient); ALS, national DEMs for broad-scale screening	Strongly resolution-dependent; fine-scale steep segments lost at coarse grids; processing choices (ground classification/interpolation) affect local extremes	ALS and public NASA DEM are efficient at large scales; SfM performance depends on ground visibility and can be sensitive to canopy/vegetation; higher-resolution products increase processing, storage, and turnaround time

7.2 Implications

7.2.1 Methodological implications: scale, constructs, and comparability

One of the main findings across this research is that mapped terrain attributes are strongly scale- and support-dependent. Traditional manual classification typically summarises terrain attributes at stand- or polygon-scale, which can obscure fine-scale variability. In contrast, remote sensing products show that derived attributes, such as roughness indices and slope gradients, vary with the spatial resolution and the neighbourhood over which they are computed. Chapter 5 demonstrated that coarsening the analysis grid from 0.5 m to 30 m systematically smoothed out micro-topography, substantially reducing the representation and extent of the steepest and roughest features.

Beyond grid resolution, DTM-derived terrain metrics are also sensitive to point-cloud processing choices, particularly ground classification and interpolation under canopy, because these steps determine which near-ground features are retained or smoothed. As a result, reproducible terrain products require documenting not only resolution but also sensor type, point density, filtering/classification settings, and the terrain-metric definition and parameterisation (e.g., neighbourhood window, detrending).

This has important methodological implications. It suggests that slope and roughness products are not unique, and they depend on the chosen spatial support, metric definition, and processing workflow. Instead, the appropriate resolution depends on the decision question. For strategic planning (e.g., road network layout), coarse-resolution ALS products at metre-scale resolution (e.g., 2–10 m) may be sufficient and computationally efficient. For tactical routing, where short steep banks or local obstacles determine trafficability, higher-resolution data (< 1 m) from UAVs or TLS may be required. Consequently, terrain classification protocols should move away from fixed labels, such as ”slope class 3”, toward scale-explicit metrics that state the resolution at which the attribute was derived. Accordingly, terrain layers should be accompanied by metadata that specifies spatial support (grid resolution and aggregation window), processing workflow, and an uncertainty indicator (e.g., expected error range or confidence class), so that planners can judge whether differences between areas reflect real terrain variation or artefacts of scale and processing.

A further challenge is comparability with established forestry TCS products. TCS frameworks were developed for plot- or polygon-based assessments in which slope, roughness, and ground conditions are summarised into discrete classes using field observations or aggregated counting rules. Remote sensing, however, produces spatially continuous raster layers at resolutions often much finer than the original TCS mapping units, making it generally difficult to apply TCS rules directly at the pixel level without reformulation. This incompatibility is most apparent for roughness and slope. In TCS, roughness is commonly ex-

pressed as obstacle frequency per unit area [1], whereas raster products represent either local obstacle density (when obstacles can be detected) or continuous surface variability. Likewise, stand-level slope classes derived from aggregated statistics can under-represent short but operationally critical steep segments. Applying TCS thresholds directly to pixels can therefore lead to different class assignments and reduced comparability with legacy terrain maps. Importantly, the correlation between different roughness layers (e.g., obstacle density and TRI) does not imply construct equivalence, because the layers may respond to distinct physical features and processing steps.

To maintain conceptual continuity while exploiting the advantages of raster data, a translation rather than a replacement approach is required. In practice, this involves deriving terrain attributes at a resolution matched to the planning task and then aggregating raster values back to decision-relevant units using explicit, documented rules. For example, obstacle-based roughness can be represented as the local obstacle density within a moving window, approximating the spatial support of the original TCS assessment. Under this framework, raster terrain products are used to derive familiar TCS-style classes in a more spatially explicit way, rather than replacing existing classification systems. Similarly, pixel-based slope can be summarised at the polygon level using explicitly defined statistics (e.g., proportions above an operational threshold or quantiles) rather than relying solely on the mean slope.

7.2.2 Operational implications for harvesting planning

The results of this work present several practical implications for the use of terrain information in mechanised harvesting planning. First, in soft-soil settings (including peatland contexts), soil-condition assessment can be refined beyond stand-level soil maps by combining wetness indicators with spatial proxies linked to vegetation structure (e.g., stump/-root influence). This can improve within-stand differentiation of relative deformation susceptibility and support extraction planning in conditions similar to those studied here.

Second, roughness products should be selected and interpreted according to the construct required by the decision task. In open or disturbed sites with high ground visibility, obstacle-based mapping from UAV photogrammetry (Chapter 3) can represent TCS-style roughness and achieved high detection accuracy in this study. Under closed canopy, where airborne visibility is limited, near-ground obstacle mapping is more appropriately derived from point-cloud–native approaches (e.g., TLS segmentation; Chapter 4), whereas DTM-derived indices (e.g., TRI) provide a practical description of continuous surface variability. Because these products represent different constructs, obstacle-density layers and DTM-based ruggedness indices should not be treated as interchangeable.

Third, remote-sensing–derived slope layers add value by representing the spatial distribution of steep segments rather than summarising slope as a single stand descriptor. Because the mapped extent of steep classes depends on grid resolution and ground-modelling choices, slope-based exclusion or system-allocation rules should be accompanied by the

source DTM, spatial support, and key processing settings.

Finally, across all pillars, the results reinforce the need to specify terrain products from the planning question backwards. Rather than selecting sensors based solely on availability, practitioners should define the decision scale (strategic, tactical, operational), the features that must be represented (e.g., local steep segments or obstacles), and the acceptable uncertainty, and then select data sources and resolutions that are feasible within those constraints. However, the findings should be interpreted in light of the scope limitations that affect transferability across sites, sensors, and operating conditions (e.g., canopy cover, soil moisture state, and processing choices).

7.2.3 FEASIBILITY IMPLICATIONS: CHOOSING SENSORS UNDER CONSTRAINTS

The operational value of remote sensing technologies in forest operations, as highlighted in Chapter 6, is determined not only by technical performance but also by practical constraints. These include survey efficiency, computational requirements for processing, data storage and management, cost considerations, and field logistics.

A critical challenge associated with rapid technological advancement is the substantial increase in data volumes. While higher-resolution sensors can provide more detailed information, the resulting datasets can become a decisive constraint [225, 260, 261]. In practice, as point density and survey extent increase, processing and storage requirements also increase, which can limit the suitability of some approaches for routine operational use. This creates a trade-off between the level of detail captured and the feasibility of its application within operational time and resource constraints.

The multi-sensor comparison in Chapter 5, together with the feasibility synthesis in Chapter 6, indicates that no single platform is optimal across all scenarios. Rather, platforms occupy different positions in the trade-off space:

- **ALS** is efficient for strategic screening and large-area mapping.
- **UAV-based systems** (LiDAR and photogrammetry) support stand- to operational-unit mapping where higher spatial detail is needed, but typically increase processing effort and data-management demands.
- **Terrestrial/handheld scanners** are most justified for targeted, plot-level or high-risk assessments where near-ground detail is required and higher effort is acceptable.

Therefore, the choice of technology should align with the decision context, balancing the required spatial detail against practical constraints such as turnaround time, processing capacity, and data management [261]. As remote sensing systems generate ever-larger datasets, managing these trade-offs will become increasingly important for practical operational use.

To reduce these bottlenecks, broader operational use will likely require a shift in focus from data acquisition alone to more efficient data management and processing. The current reliance on manual post-processing can slow operational adoption when data volumes are high [262]. A practical pathway is the development of automated processing pipelines (including cloud- or edge-computing approaches) that translate raw sensor data into decision-ready products with reduced user effort [261].

Equally critical is demonstrating economic viability. Future research should move beyond technical feasibility to quantify return on investment using transparent, scenario-based analyses calibrated to specific forest types and operational contexts. For example, on steep terrain, higher-resolution terrain products can help planners identify local steep segments and choose routes that avoid problematic sections. This could reduce fuel use, machine strain, or time spent replanning, but potential savings need to be evaluated against the costs of data acquisition and processing. Therefore, developing and validating such cost-benefit models would provide managers with clearer evidence to justify technology investments and support the transition from pilot applications to routine operational use. In this context, the economic justification for remote-sensing technologies also depends on the time horizon over which terrain data are used and the scale of production they support.

From an operational perspective, the usefulness of higher-resolution terrain sensing depends mainly on how the data will be used and the scale of the harvesting operation. In the *short term*, existing ALS or publicly available DEM products can already support broad terrain screening and the identification of basic constraints at relatively low cost per hectare. Because these datasets often cover large areas, they can also be reused across multiple planning cycles. In the *medium term*, the main opportunity lies in reducing costs through more automated data processing and by linking terrain information with operational data streams (e.g., machine telemetry) to improve calibration and update terrain assessments over time. In the *long term*, terrain information may increasingly support machine-aware systems, where prior terrain maps are combined with onboard sensors to update traversability estimates during operation. However, such developments depend on the availability of standardised workflows that can efficiently translate raw data into decision-ready products [261]. In practice, higher-cost sensing approaches (e.g., UAV LiDAR or terrestrial scanning) are most likely to be justified in operations with high harvested volumes, complex terrain conditions such as steep or wet sites, or situations where terrain data can be reused across several planning cycles and the acquisition costs can be distributed over multiple operations. However, these expectations remain largely conceptual and require empirical evaluation in operational forestry contexts to determine under which conditions the additional sensing costs are consistently justified.

In relation to RO4 and RQ4 (Section 1), the findings show that operational feasibility is governed by a trade-off between required spatial detail and constraints on survey effort, data volume, processing/turnaround time, cost, and field logistics (Chapter 6). Consequently, platform choice should be made from the decision scale backwards: ALS is best

suited to large-area screening, UAV-based approaches are suitable for stand- and operation-level planning where higher resolution is needed, and processing capacity is available, and TLS/handheld approaches are most justified for targeted high-risk assessments where near-ground detail is required.

7.3 Validation against operational outcomes

Validation in the context of mechanised harvesting can be understood as a hierarchy of evidence. Terrain products may be validated against (a) field observations of terrain features (e.g., obstacle counts), (b) measured site responses to traffic (e.g., rut depth), or (c) operational outcomes, such as realised route choices, mobility constraints, productivity, fuel consumption, or the occurrence of soil disturbance and damage. Linking terrain layers to operational outcomes is challenging because performance is influenced not only by terrain but also by operator decisions, machine configuration, time-varying soil moisture and seasonality, and the limited positional accuracy of many onboard GNSS systems under canopy.

Accordingly, the studies in this dissertation provide targeted field validation, comparative evaluation, and conceptual synthesis, rather than end-to-end validation against operational outcomes for all derived products. Table 7.3 summarises which terrain products were directly validated, which were evaluated through comparative analyses, and which remain conceptual.

7.4 Limitations and uncertainty

Several limitations should be considered when interpreting the results. First, ground conditions were assessed primarily through observed deformation responses and spatial proxies (e.g., stump/root influence, wetness indicators, and machine variables), supported by soil shear-strength measurements. This limits inference about absolute bearing capacity. Instead, the outputs should be interpreted as relative susceptibility maps. In addition, the stump/root influence model assumes simplified spatial behaviour and does not represent anisotropic root architectures or their variation with species, stand age, and site conditions, as discussed in previous sections.

Second, roughness outcomes depend strongly on visibility and processing choices. Obstacle-based mapping was developed under conditions of high ground visibility, whereas under a closed canopy, near-ground objects may be obscured, and ground filtering/interpolation may suppress discrete obstacles. As a result, transferring obstacle-based approaches to dense-canopy environments is uncertain and requires additional validation and, potentially, alternative sensing and processing strategies. The exploratory TLS-based work demonstrated the potential to detect and classify near-ground objects directly from point clouds. Future studies should therefore further investigate point-cloud-native ap-

Table 7.3: Scope and strength of validation for key terrain products developed in this dissertation. "Field validation" refers to comparison against measured site features or site responses; "comparative evaluation" to cross-sensor or multi-resolution analyses; "conceptual" to synthesis not yet tested as an end-to-end predictive model.

Terrain product	Evidence, type	Validated/evaluated against	Key gaps, limitations
Rut depth prediction model using stump/root reinforcement proxies	Field validation (mechanical response)	Observed rut depth; supporting mechanical variables (e.g., soil strength, machine/loading descriptors)	Not yet linked to productivity, fuel use, route choices, or long-term soil/stand impacts; root reinforcement simplified and species-agnostic, with no explicit root architecture or anisotropy
Obstacle-based roughness mapping model (UAV; open terrain)	Field validation + comparative evaluation	Field obstacle observations (counts/density); comparison with DTM-based indices (TRI/VRM/AR)	Not yet validated against mobility, productivity, fuel use, or operator routing; transfer depends on ground visibility
DTM-derived slope and ruggedness	Comparative evaluation (cross-sensor / multi-resolution)	DTM co-registration checks; cross-sensor and multi-resolution comparisons of slope and ruggedness	Not yet linked to operational outcomes; sensitivity to grid resolution and point-cloud processing, especially under canopy
Conceptual trafficability framework (synthesis across pillars)	Conceptual synthesis	Integration of the three pillars into a planning-oriented framework (not implemented as an end-to-end predictive model)	End-to-end performance not tested; requires explicit decision rules/thresholds, local calibration, and operational datasets (e.g., machine tracks, telemetry, disturbance observations)

proaches for roughness characterisation under canopy, rather than relying solely on DTM-derived raster metrics.

Third, slope and ruggedness estimates are sensitive to point-cloud ground classification and interpolation in closed-canopy or otherwise vegetated terrain. Differences in sensor characteristics and processing settings can propagate into derived slope classes, particularly in complex terrain where local extremes are operationally most relevant.

Fourth, the generalisability of data-driven models is constrained by site-specific training and validation data (e.g., the rut-depth prediction model and the roughness classification framework). While the workflows are transferable in principle, model parameters, thresholds, and expected performance should be re-evaluated when applying the methods to different forest types, terrain conditions, or sensor configurations.

Finally, the terrain products represent conditions at the time of data acquisition. Because soil trafficability and deformation risk can vary substantially with soil moisture and seasonal conditions, the mapped susceptibility patterns should not be interpreted as tem-

porally invariant. This limitation is especially relevant when extrapolating to wetter periods, freeze–thaw conditions, or seasons with different understorey and ground-surface visibility.

Taken together, the results provide the most substantial evidence for (a) the feasibility of producing spatially explicit terrain layers from remote sensing in the study context and (b) the relative behaviour of key terrain metrics across scales and processing choices, including how resolution and ground modelling influence derived slope, ruggedness, and roughness representations. The mapped outputs should therefore be interpreted primarily as decision-support layers that describe relative susceptibility and spatial variability, rather than as absolute estimates of bearing capacity or guaranteed operational performance.

Conclusions become more conditional when transferring the approaches to settings that differ materially from the study conditions (e.g., dense-canopy environments with low near-ground visibility, different forest types, terrain conditions, or sensor configurations) and when interpreting terrain layers as predictors of productivity, fuel/effort, safety, or soil-impact outcomes without route-linked operational validation data. In these cases, thresholds and model parameters should be treated as provisional until recalibrated and tested against observed machine-route outcomes.

To reduce uncertainty in how mapped terrain layers translate into machine performance and soil impact (without requiring complete system optimisation), a minimal operational validation dataset should link pre-harvest terrain products to outcomes observed along actual machine routes. A feasible minimum could include:

- **Machine tracks and activity:** Time-stamped GNSS trajectories for ground-based harvesting machinery (e.g., harvesters, forwarders, skidders), preferably segmented into basic operational states (e.g., travelling empty, travelling loaded, loading/unloading).
- **Production and load:** Payload or load-state indicators (where available) and simple time-based productivity measures (e.g., travel time, productive machine hours).
- **Effort and mobility:** Fuel consumption or engine-load proxies (e.g., RPM/load, hydraulic pressure) and, if available, mobility proxies such as wheel slip or speed reductions.
- **Soil disturbance:** Post-harvest rut observations along selected track segments (field or high-resolution surveys), reported as maximum rut depth and/or the proportion exceeding a rut threshold.

Validation should also account for GNSS uncertainty under canopy, scale mismatch between map resolution and machine footprint, and time-varying ground conditions by matching terrain metrics to buffered track segments and including moisture-state and

seasonal covariates. Analyses can then test whether predicted soil-risk relates to rut exceedance, roughness relates to mobility or route deviations, and slope metrics relate to effort (fuel or engine load) and productivity components.

7.5 FUTURE RESEARCH DIRECTIONS

A consistent theme across the empirical studies is that remote sensing can quantify terrain attributes relevant to mechanised forest operations. However, that operational uptake requires additional steps beyond producing stand-alone terrain layers. Future research should prioritise integrating the three terrain pillars into decision-ready products, validating and calibrating these products against machine behaviour and soil impacts, and extending terrain assessment from static pre-processed maps toward dynamic and, ultimately, autonomous operation.

7.5.1 INTEGRATION OF THE THREE PILLARS INTO ACCESSIBLE DECISION TOOLS

A key barrier to industry adoption is that remote sensing outputs are often delivered as separate layers for slope, roughness, and ground conditions, requiring practitioners to synthesise the information manually. While Nordic GIS-based trafficability tools (e.g., *Metsään.fi*, Harvester Seasons) have integrated soil type, wetness, and topography into simple colour-coded risk maps [35, 263], these solutions are typically coarse-resolution, region-specific, and not explicitly aligned with the three-pillar framework used in this research. Similarly, advanced traversability models, such as those of Wallin et al. [149], predict continuous trafficability from terrain data but rely on synthetic simulations rather than on field-validated, high-resolution inputs.

Future research should therefore prioritise developing integrated decision-support products that combine the three pillars—ground conditions (Chapter 2), surface roughness (Chapters 3, 4), and slope (Chapter 5), into a single, scalable representation for planning. A practical implementation would require (a) aligning the three input layers to a common spatial support appropriate to the decision scale, (b) transforming each layer into a comparable constraint or risk score through class-based re-scaling, and (c) combining the scores using transparent weighting schemes. The resulting product could be expressed either as a continuous score (representing increasing mobility constraints or disturbance risk) or as a small number of decision classes (e.g., *suitable*, *constrained*, *unsuitable*) defined by clearly stated thresholds.

To remain operationally relevant, the index could be machine-agnostic in its base form, with optional calibration parameters for equipment configuration (e.g., tyre vs tracked undercarriages, variable payloads, and assisted vs non-assisted harvesting) and seasonal state (e.g., wet vs frozen conditions). This requires linking index outputs to measurable operational outcomes, such as rutting occurrence, wheel slip, vibration, speed reductions,

or fuel consumption, and updating the ground-condition component with time-varying wetness or moisture information where available. Finally, the index could be delivered through user-centred interfaces (dashboards, web services, or in-cab displays) that expose both the final classes and the underlying drivers (e.g., "high risk due to wetness" versus "high risk due to slope"), thereby shifting the deliverable from multiple layers to interpret toward decision-ready products that can be acted on in planning and operations. Related approaches are being operationalised in Scandinavian industry through operational trafficability and wetness products, such as Arbonaut's forest trafficability maps and the Finnish and Swedish wet area map services, which provide planners with simplified, map-based indicators of rutting risk for day-to-day decision-making [264, 265].

7.5.2 Validation and model calibration against operational outcomes

However, developing such integrated, decision-ready products requires calibration against operational outcomes. It demands robust calibration against operational outcomes. While emerging research has begun to correlate machine CAN-bus signals, such as fuel consumption, wheel slip, and engine load, with trafficability parameters [**?**], these methods are not yet widely used to systematically calibrate remote-sensing-derived terrain indices.

Future studies could link high-resolution terrain metrics to machine performance by spatially correlating terrain layers with georeferenced CAN bus telemetry. Treating the forest machine as an observation source, researchers could quantify how terrain features, such as obstacle height or slope gradient, relate to changes in fuel consumption, wheel slip, vibration, or travel speed. This validation loop would enable calibration of decision-support products, moving from fixed thresholds to empirically supported relationships.

However, implementation faces practical challenges. Manufacturers often restrict access to proprietary machine data (e.g., CAN bus), and standard onboard GNSS can be insufficiently precise to align telemetry with high-resolution terrain models. Successfully implementing this validation loop will therefore require close collaboration with equipment manufacturers to access data streams and to integrate high-precision positioning systems (e.g., RTK-GNSS) into standard harvesting workflows.

7.5.3 From static maps to autonomous perception

The high-resolution terrain characterisation developed in this work may also be relevant to longer-term transitions toward increased automation in forest operations [266]. Automated machines typically require explicit, quantitative representations of traversability to support robust operation in complex environments. This motivation is reflected in the development of high-resolution multimodal datasets, such as FORWARD [267], which capture interactions between heavy machinery and rough terrain for training and evaluation of perception algorithms. As sensing shifts from overhead mapping toward onboard per-

ception using SLAM, stereo vision, and LiDAR [268, 269], terrain metrics may increasingly need to support near-real-time interpretation in addition to planning-stage products.

A key methodological challenge is developing workflows that integrate a prior terrain model (or map) with onboard perception, enabling traversability estimates to be updated as the machine moves. Prototype systems reported in the literature indicate the feasibility of harvester-mounted laser scanning for SLAM-based mapping and path tracking in forest environments [270, 271, 272]. Related work on legged robots (e.g., ANYmal) using graph-based SLAM further demonstrates that autonomous navigation and mapping under a canopy are possible in principle [269]. Together, these developments suggest a pathway to integrate pre-harvest terrain layers with onboard perception to enable machine reasoning about local traversability.

7.5.4 Potential future studies

Future research building on this work can retain the three-pillar structure used throughout the thesis while strengthening links to operational outcomes and explicitly accounting for data-feasibility constraints.

Soil conditions and deformation response.

The rut-depth modelling approach developed can be extended in three directions:

- **Outcome-linked validation:** Evaluate whether predicted rutting susceptibility relates to observed disturbance along operational extraction routes by combining machine trajectories and available telemetry (e.g., travel speed, wheel slip where available, engine-load proxies) with post-harvest rut measurements.
- **Transferability testing:** Test model performance across a broader range of conditions, including different soil types and moisture regimes, seasonal states (e.g., unfrozen vs frozen), and machine configurations (e.g., axle loads; wheeled vs tracked; assisted vs non-assisted systems).
- **Refinement of stump/root proxies:** Develop species- and site-sensitive parameterisations that better reflect variation in root architecture and mechanical properties, supported where feasible by targeted field measurements, TLS-based structural mapping, or integration with existing root-reinforcement formulations.

Surface roughness and obstacle representation.

Future work on roughness can focus on improving transferability beyond the experimental post-fire setting:

- **Cross-site evaluation:** Quantify how obstacle-detection and roughness mapping performance varies across disturbance histories, vegetation structures, and obstacle types, including sensitivity to spatial resolution and sensor characteristics.

- **Raster-compatible decision logic:** Reformulate plot-/polygon-based roughness thresholds into explicit raster aggregation and decision rules (e.g., window sizes and density thresholds), ideally refined through expert elicitation and testing with practitioners.
- **Point-cloud–native obstacle mapping:** Further develop and validate point-cloud–native workflows (e.g., TLS/mobile LiDAR) for near-ground obstacle detection where DTM/DSM products suppress or smooth critical features.
- **Construct-appropriate benchmarking:** Define and test when obstacle-based roughness products versus continuous ruggedness indices are appropriate for specific planning decisions, rather than treating them as interchangeable.

Steepness and terrain complexity.

Future studies can strengthen generalisability and separate sensor effects from processing effects:

- **Broader site testing:** Replicate analyses across a wider range of terrain types and forest conditions to test the robustness of observed sensor- and scale-dependencies.
- **Controlled processing experiments:** Systematically vary ground-point density, interpolation method, and ground filtering/classification parameters on common point-cloud inputs to isolate processing-induced effects.
- **Planning-scale specifications:** Derive minimum data specifications (sensor type, point density, grid resolution, processing settings) needed to represent slope classes and terrain-complexity metrics reliably at different planning scales.

Cross-cutting validation and data feasibility.

Across all pillars, a priority is outcome-linked validation using operational datasets, supported by feasible processing pipelines:

- **Co-located outcome datasets:** Collect co-registered terrain products, machine trajectories/telemetry, and post-harvest disturbance observations over the same sites and time windows to enable direct terrain–outcome tests.
- **Feasible data pipelines:** Design workflows around data volume, processing effort, and turnaround time, including automated preprocessing and selective data retention, to fit operational planning cycles.

Chapter 8

Summary and conclusions

This dissertation evaluated the potential of remote sensing-based terrain analysis to quantify terrain attributes relevant to mechanised timber-harvesting planning and to assess how reliably these attributes can be derived across contrasting forest environments, sensors, and spatial scales. The work was organised around the three terrain components commonly represented in forestry terrain classification systems (TCS): soil conditions, surface roughness, and slope.

The results show that remote sensing enables a shift from stand- or polygon-level terrain descriptors toward spatially explicit terrain layers (raster- and, in some cases, point-cloud–based) that reveal within-stand variability relevant to routing, exclusion mapping, and mitigation planning. At the same time, the work shows that terrain metrics are not scale-invariant: the mapped representation depends on spatial support (e.g. resolution), sensor characteristics, and processing choices. Consequently, terrain layers must be generated and interpreted with respect to the planning task and the terrain construct being represented.

- **Soil conditions (RO1/RQ1)**

  In soft-soil conditions, stump distribution and associated root reinforcement can act as a quantifiable proxy for within-site variability in deformation response under repeated machine traffic. Within the study conditions, models that combined terrain-derived predictors with operational variables achieved $R^2 = 0.69$–0.85 for observed rut depth across machine passes and trail configurations, supporting relative rutting-susceptibility mapping at extraction-trail scales. The findings also delimit transferability: application beyond the study conditions requires recalibration and validation across differing soil types, moisture states/seasons, and machine configurations, and the current root influence formulation should be interpreted as a simplified representation of reinforcement rather than an explicit root-architecture model.

- **Surface roughness (RO2/RQ2)**

  Obstacle-based roughness, as defined in forestry TCS, can be operationalised from high-resolution UAV data where near-ground visibility is sufficient. In the open-visibility post-fire setting, a deep-learning segmentation workflow detected discrete obstacles with 95.6% accuracy (86/90) relative to field-marked obstacles. It enabled spatial roughness

mapping at 3 × 3 m support using local obstacle density. Comparisons with DTM-derived ruggedness metrics (TRI, VRM, AR) showed that some indices reproduce broad spatial trends, but they represent a different construct (continuous surface variability rather than discrete obstacle frequency). Therefore, correlation with DTM indices should be interpreted as partial alignment rather than construct equivalence. In closed-canopy or dense near-ground vegetation, where obstacle visibility is limited, obstacle-based roughness cannot be assumed to be observable in airborne products. In such settings, point-cloud-native approaches (e.g., TLS-based segmentation) provide a means to characterise discrete obstacles, whereas DTM metrics remain suitable for describing broader surface variability.

- **Steepness and terrain complexity (RO3/RQ3)**

  Steepness and related terrain-complexity outcomes derived from DTMs are strongly sensitive to spatial resolution and to ground-surface modelling under canopy. Fine-scale steep segments and local extremes relevant to operational safety and detailed mitigation are progressively smoothed at coarser grid sizes, reducing the mapped extent of steep classes and altering slope-category outcomes. Within the comparative evaluation, slope estimates were comparatively stable at intermediate spatial supports (approximately 2–4 m) when ground-point densities were sufficient (on the order of 5–10 pts m^{-2}). However, the optimal specification depends on terrain complexity and the planning task. For broad-scale screening, ALS DTMs and publicly available DEM products provided fit-for-purpose slope information for baseline constraint mapping. Ruggedness indices such as TRI can complement slope layers, but they require cautious interpretation because they can be slope-coupled at different scales and terrain types.

- **Feasibility and operational fit (RO4/RQ4)**

  This dissertation also indicates that operational uptake depends on more than technical accuracy: survey effort, data volume, processing demands, and turnaround time determine whether a workflow is feasible within the planning stage. No single platform is optimal across all use cases. Instead, platform choice should be made from the decision scale backwards: ALS supports efficient large-area screening, UAV-based systems provide stand- and operation-level detail where finer spatial support is required, and terrestrial/handheld approaches are most justified for targeted, high-risk assessments where near-ground detail is essential.

Aligned with RO1–RO4 and RQ1–RQ4 (Chapter 1), the main contributions are:

- A stump/root influence proxy and rut-depth prediction framework that supports relative rutting-susceptibility mapping under repeated traffic in soft soils, with documented transferability constraints.

- An automated obstacle-based roughness workflow for open terrain using UAV photogrammetry and deep-learning segmentation, together with an explicit construct boundary between obstacle-based roughness and DTM-derived ruggedness metrics. Exploratory results also indicate the utility of point-cloud–native TLS approaches where airborne visibility is limited.
- Quantified evidence of how sensor type, point density, spatial support, and processing choices influence DTM-derived slope classes and terrain-complexity metrics under canopy, supporting scale-aware use of slope layers as constraint products.
- A cross-platform synthesis that links sensing options to decision scale and feasibility constraints to support evidence-based selection of terrain products for planning.

In conclusion, remote sensing-based terrain analysis can improve the objectivity and spatial specificity of terrain information used in mechanised harvesting planning by representing soil-condition proxies, roughness constructs, and slope constraints as explicit spatial layers rather than stand-level descriptors. Effective use, however, requires explicit alignment between (a) the planning objective, (b) the terrain construct being represented, and (c) the spatial support and processing workflow used to derive the product. The thesis provides a methodological foundation for integrating remote-sensing terrain layers into planning. It defines the key requirements for future work that links these layers to operational outcomes and integrates the three pillars into decision-ready products.

Bibliography

[1] Erasmus, D., 1994. National terrain classification system for forestry, version 1.0. Technical Report 117/94, Institute for Commercial Forestry Research, Pietermaritzburg, South Africa.

[2] International Society for Terrain-Vehicle Systems, 1977. International society for terrain-vehicle systems standards. *Journal of Terramechanics*, 14(3):153–182. doi: 10.1016/0022-4898(77)90013-1.

[3] He, R., Sandu, C., Mousavi, H., Shenvi, M. N., Braun, K., Kruger, R., and Els, P. S., 2020. Updated Standards of the International Society for Terrain-Vehicle Systems. *Journal of Terramechanics*, 91:185–231.

[4] Food and Agriculture Organisation of the United Nations, 2024. The State of the World's Forests 2024: Forest-sector innovations towards a more sustainable future. Technical report, Food and Agriculture Organization of the United Nations, Rome. URL https://doi.org/10.4060/cd1211en.

[5] Food and Agriculture Organization of the United Nations, 2024. Global forest products facts and figures 2023. URL https://doi.org/10.4060/cd3650en. FAO Statistics: industrial roundwood, sawnwood, wood-based panels, fibre furnish, paper and paperboard, wood fuel, charcoal and pellets.

[6] Eurostat, 2023. Wood products – production and trade. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Wood_products_-_production_and_trade. Statistics Explained, European Commission. Accessed 2025-12-16.

[7] Moskalik, T., Borz, S. A., Dvořák, J., Ferencik, M., Glushkov, S., Muiste, P., Lazdiņš, A., and Styranivsky, O., 2017. Timber harvesting methods in Eastern European countries: A review. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 38(2):231–241.

[8] Lundbäck, M., Häggström, C., and Nordfjell, T., 2021. Worldwide trends in methods for harvesting and extracting industrial roundwood. *International Journal of Forest Engineering*, 32(3):202–215.

[9] Erber, G. and Spinelli, R., 2020. Timber extraction by cable yarding on flat and wet terrain: a survey of cable yarder manufacturer's experience. *Silva Fennica*, 54(2).

[10] Eberhard, B., Trailovic, Z., Magagnotti, N., and Spinelli, R., 2025. A GIS-Based Decision Support Model (DSM) for Harvesting System Selection on Steep Terrain: Integrating Operational and Silvicultural Criteria. *Forests*, 16(5):854.

[11] Visser, R. and Berkett, H., 2015. Effect of terrain steepness on machine slope when harvesting. *International Journal of Forest Engineering*, 26(1):1–9.

[12] Mologni, O., Grigolato, S., Cavalli, R., et al., 2016. Harvesting systems for steep terrain in the Italian Alps: state of the art and future prospects. *Contemporary Engineering Sciences*, 9(25):1229–1242.

[13] Proto, A. R. and Zimbalatti, G., 2016. Firewood cable extraction in the southern Mediterranean area of Italy. *Forest Science and Technology*, 12(1):16–23. ISSN 2158-0103.

[14] Lundbäck, M., Persson, H., Häggström, C., and Nordfjell, T., 2021. Global analysis of

the slope of forest land. *Forestry: An International Journal of Forest Research*, 94(1): 54–69.

[15] Picchio, R., Latterini, F., and Venanzi, R., 2026. GHG Emissions from Forest Operations in Mediterranean Chestnut Coppices. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 47(1):55–68.

[16] Cambi, M., Certini, G., Neri, F., and Marchi, E., 2015. The impact of heavy traffic on forest soils: A review. *Forest Ecology and Management*, 338:124–138.

[17] Hertz, E., Guriev, A., Druzyanova, V., Revyako, S., Markov, O., Perfiliev, P., and Grigorev, I., 2024. Impact Assessment of Different Propulsion Systems in Forestry Machinery on Soil Properties. *Mathematical Modelling of Engineering Problems*, 11(1):133–140. doi: 10.18280/mmep.110114.

[18] Liepiņš, K., Lazdiņš, A., Liepiņš, J., and Prindulis, U., 2015. Productivity and cost-Effectiveness of Mechanized and Motor-Manual Harvesting of Grey Alder (*Alnus incana* (L.) Moench): A Case Study in Latvia. *Small-scale Forestry*, 14(4):493–506.

[19] FAO, I. and UNECE, 2023. Occupational safety and health in forest work. Technical report, Food and Agriculture Organization of the United Nations (FAO); International Labour Organization (ILO); United Nations Economic Commission for Europe (UNECE), Geneva. URL https://unece.org/sites/default/files/2023-07/Forest%20work%20safety%20FAO-ILO-UNECE%20report.pdf. Accessed: 2025-12-16.

[20] Marušiak, M. and Neruda, J., 2018. Dynamic soil pressures Caused by Travelling Forest Machines. *Croatian Journal of Forest Engineering*, 39(2):233–245. ISSN 1845-5719. Original scientific paper.

[21] Pirnazarov, A. U., 2015. *Forestry machine and soil interaction for sustainable forestry*. Doctoral thesis, KTH Royal Institute of Technology, Stockholm, Sweden.

[22] Prinz, R., Mola-Yudego, B., Ala-Ilomäki, J., Väätäinen, K., Lindeman, H., Talbot, B., and Routa, J., 2023. Soil, driving speed and driving intensity affect fuel consumption of forwarders. *Croatian Journal of Forest Engineering*, 44(1):31–43.

[23] Heppelmann, J. B., Talbot, B., Antón-Fernández, C., and Astrup, R., 2022. Depth-to-water maps as predictors of rut severity in fully mechanized harvesting operations. *International Journal of Forest Engineering*, 33(2):108–118.

[24] Salmivaara, A., Holmström, E., Kulju, S., Ala-Ilomäki, J., Virjonen, P., Nevalainen, P., Heikkonen, J., and Launiainen, S., 2024. High-resolution harvester data for estimating rolling resistance and forest trafficability. *European Journal of Forest Research*, 143(6):1641–1656.

[25] Jang, M.-K., Hwang, S.-J., Kim, J.-H., and Nam, J.-S., 2022. Overturning and rollover characteristics of a tractor through dynamic simulations: Effect of slope angle and obstacles on a hard surface. *Biosystems Engineering*, 219:11–24. ISSN 1537-5110.

[26] Vega-Nieva, D. J., Murphy, P. N. C., Castonguay, M., Ogilvie, J., and Arp, P. A., 2009. A modular terrain model for daily variations in machine-specific forest soil trafficability. *Canadian Journal of Soil Science*, 89(1):93–109.

[27] Verkerk, P. J., Hassegawa, M., Van Brusselen, J., Cramm, M., Chen, X., Maximo, Y. I., Koç, M., Lovrić, M., and Tegegne, Y. T., 2022. *The role of forest products in the global bioeconomy: Enabling substitution by wood-based products and contributing to the Sustainable Development Goals*. Food and Agriculture Organization of the United Nations (FAO), Rome. ISBN 978-92-5-135151-2. doi: 10.4060/cb7274en. URL https://doi.org/10.4060/cb7274en. FAO report.

[28] Cadei, A., Mologni, O., Röser, D., Cavalli, R., and Grigolato, S., 2020. Forwarder Productivity in Salvage Logging Operations in Difficult Terrain. *Forests*, 11(3):341.

[29] Ackerman, P., Martin, C., Brewer, J., and Ackerman, S., 2018. Effect of slope on pro-

ductivity and cost of Eucalyptus pulpwood harvesting using single-grip purpose-built and excavator-based harvesters. *International Journal of Forest Engineering*, 29(2):74–82.

[30] Potočnik, I., Pentek, T., and Poje, A., 2009. Severity Analysis of Accidents in Forest Operations. *Croatian Journal of Forest Engineering*, 30(2):171–184. URL https://hrcak.srce.hr/47892.

[31] Poje, A., Grigolato, S., and Potočnik, I., 2019. Operator Exposure to Noise and Whole-Body Vibration in a Fully Mechanised CTL Forest Harvesting System in Karst Terrain. *Croatian Journal of Forest Engineering*, 40(1):139–150.

[32] Cavalli, R. and Amishev, D., 2019. Steep terrain forest operations–challenges, technology development, current implementation, and future opportunities. *International Journal of Forest Engineering*, 30(3):175–181.

[33] Province of British Columbia, 1996. Forest Practices Code of British Columbia Act. https://www.bclaws.gov.bc.ca/civix/document/id/complete/statreg/96159_01. [RSBC 1996] Chapter 159.

[34] Arendarczyk, B., Rabin, S., Bampoh, D., Arneth, A., Rounsevell, M., and Alexander, P., 2025. Response of Global Forest Management to Changes in Wood Demand. *Global Change Biology*, 31(11):e70573.

[35] Kankare, V., Luoma, V., Saarinen, N., Peuhkurinen, J., Holopainen, M., and Vastaranta, M., 2019. Assessing feasibility of the forest trafficability map for avoiding rutting—a case study. *Silva Fennica*, 53(3).

[36] Ulvdal, P., Öhman, K., Eriksson, L. O., Wästerlund, D. S., and Lämås, T., 2023. Handling uncertainties in forest information: the hierarchical forest planning process and its use of information at large forest companies. *Forestry*, 96(1):62–75.

[37] Shoop, S. A., June 1993. Terrain Characterization for Trafficability. Technical Report CRREL Report 93-6, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire. Prepared for Office of the Chief of Engineers.

[38] Christian, C. S. and Stewart, G. A., 1953. *General Report on Survey of Katherine-Darwin Region, 1946*. Number 1 in Land Research Series. Commonwealth Scientific and Industrial Research Organization (CSIRO), Melbourne. ISBN 0-643-00651-6.

[39] Beckett, P. H. T., Webster, R., McNeil, G. M., and Mitchell, C. W., 1972. Terrain Evaluation by Means of a Data Bank. *The Geographical Journal*, 138(4):430–449.

[40] Haarlaa, R., 1975. The effect of terrain on the output in forest transportation of timber. *Journal of Terramechanics*, 12(2):55–94.

[41] Mellgren, P. G., 1980. Terrain Classification for Canadian Forestry. Technical report, Canadian Pulp and Paper Association.

[42] Davis, C. J. and Reisinger, T. W., 1990. Evaluating terrain for harvesting equipment selection. *Journal of Forest Engineering*, 2(1):9–16.

[43] Berg, S., 1995. *Terrängtypsschema för skogsarbete [Terrain classification scheme for forest operations]*. SkogForsk, Uppsala, Sweden. ISBN 91-7614-035-0. URL https://www.skogforsk.se/contentassets/dd0282d1b35c4fe88f210b088f02b486/terrangtypschema.pdf.

[44] Rowan, A. A., 1995. Terrain Classification. Technical Note 16/95, British Forestry Commission, Forest Research.

[45] Terlesk, C. J., May 1983. A proposed terrain classification for harvesting in New Zealand. *New Zealand Journal of Forestry*, 28(1):43–57.

[46] Bazarov, S. M., Alexsandrov, V. A., Belenky, Y. I., Gasimov, G. S., and Kalistratov, A. V., 2020. Physical and mechanical properties of forest soils: a review. In *IOP Conference Series: Materials Science and Engineering*, volume 817, page 012002. IOP Publishing.

[47] Hoffmann, S., Schönauer, M., Heppelmann, J., Asikainen, A., Cacot, E., Eberhard, B.,

Hasenauer, H., Ivanovs, J., Jaeger, D., Lazdins, A., et al., 2022. Trafficability prediction using depth-to-water maps: the status of application in northern and central European forestry. *Current Forestry Reports*, 8(1):55–71.

[48] Kaźmierowski, C., Ceglarek, J., Królewicz, S., Cierniewski, J., Jasiewicz, J., and Wyczałek, M., 2015. Soil surface roughness quantification using DEM obtained from UAV photogrammetry. *Geomorphom Geosci, Adam Mickiewicz University in Poznań–Institute of Geoecology and Geoinformation*. doi: 10.

[49] Niemi, M. T., Vastaranta, M., Vauhkonen, J., Melkas, T., and Holopainen, M., 2017. Airborne LiDAR-derived elevation data in terrain trafficability mapping. *Scandinavian Journal of Forest Research*, 32(8):762–773.

[50] Latterini, F., Venanzi, R., Tocci, D., Moschetti, F., and Picchio, R., 2020. Comparing accuracy of three remote sensing methods to evaluate soil impact related to forest operations. *Environmental Sciences Proceedings*, 3(1):59.

[51] Schönauer, M., Ågren, A. M., Katzensteiner, K., Hartsch, F., Arp, P., Drollinger, S., and Jaeger, D., 2024. Soil moisture modeling with ERA5-Land retrievals, topographic indices, and in situ measurements and its use for predicting ruts. *Hydrology and Earth System Sciences*, 28(12):2617–2633.

[52] Talbot, B., Pierzchała, M., and Astrup, R., 2017. Applications of remote and proximal sensing for improved precision in forest operations. *Croatian Journal of Forest Engineering*, 38(2):327–336.

[53] Tomljanović, K., Kolar, A., Đuka, A., Franjević, M., Jurjević, L., Matak, I., Ugarković, D., and Balenović, I., 2022. Application of UAS for Monitoring of Forest Ecosystems – A Review of Experience and Knowledge. *Croatian Journal of Forest Engineering*, 43 (2):487–504.

[54] Murphy, P. N., Ogilvie, J., Connor, K., and Arp, P. A., 2007. Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada. *Wetlands*, 27(4): 846–854.

[55] Beven, K. J. and Kirkby, M. J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological sciences journal*, 24(1):43–69.

[56] Salmivaara, A., Mäkitalo, L., Kankare, V., Väätäinen, K., and Pitkänen, S., 2020. Towards dynamic forest trafficability prediction using open spatial data, hydrological modelling and sensor technology. *Forestry*, 93(5):662–674.

[57] Sterenczak, K. and Moskalik, T., 2014. Use of LIDAR-based digital terrain model and single tree segmentation data for optimal forest skid trail network. *iForest-Biogeosciences and Forestry*, 8(5):661.

[58] Mohtashami, S., 2022. *GIS-based decision support systems to minimise soil impacts in logging operations*. Doctoral thesis, Swedish University of Agricultural Sciences, Umeå, Sweden. URL https://doi.org/10.54612/a.qq3cqbcknd.

[59] Hrza, P., Mikita, T., and Žižlavská, N., 2025. Application of Unmanned Aerial Vehicle and Airborne Light Detection and Ranging Technologies to Identifying Terrain Obstacles and Designing Access Solutions for the Interior Parts of Forest Stands. *Forests*, 16(5):729.

[60] Latterini, F., Camarretta, N., and Watt, M. S., 2025. Remote sensing for planning harvesting operations and monitoring their effects on the forest ecosystem: State of the art and future perspectives. *Forest Ecology and Management*, 597:123175.

[61] Smreček, R., 2012. Utilization of ALS data for forestry purposes. In *GI Forum*, pages 365–375.

[62] Yrttimaa, T., Matsuzaki, S., Kankare, V., Junttila, S., Saarinen, N., Kukko, A., Hyyppä, J., Miura, J., and Vastaranta, M., 2024. Assessing Forest Traversability for Autonomous

Mobile Systems Using Close-Range Airborne Laser Scanning. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 45 (1):169–182.

[63] Koren, M., Slančík, M., Suchomel, J., and Dubina, J., 2015. Use of terrestrial laser scanning to evaluate the spatial distribution of soil disturbance by skidding operations. *iForest-Biogeosciences and Forestry*, 8(3):386.

[64] Wu, Y., Zhong, S., Ma, Y., Zhang, Y., and Liu, M., 2025. Application of SLAM-based mobile laser scanning in forest inventory: Methods, progress, challenges, and perspectives. *Forests*, 16(6):920.

[65] Pierzchała, M., Giguère, P., and Astrup, R., 2018. Mapping forests using an unmanned ground vehicle with 3D LiDAR and graph-SLAM. *Computers and Electronics in Agriculture*, 145:217–225.

[66] Salmivaara, A., Miettinen, M., Finér, L., Launiainen, S., Korpunen, H., Tuominen, S., Heikkonen, J., Nevalainen, P., Sirén, M., Ala-Ilomäki, J., et al., 2018. Wheel rut measurements by forest machine-mounted LiDAR sensors–accuracy and potential for operational applications? *International Journal of Forest Engineering*, 29(1):41–52.

[67] Kim, I., Seo, J., Woo, H., and Choi, B., 2025. Assessing Rutting and Soil Compaction Caused by Wood Extraction Using Traditional and Remote Sensing Methods. *Forests*, 16(1):86.

[68] Jones, M. and Arp, P., 2019. Soil Trafficability Forecasting. *Open Journal of Forestry*, 9(4):296–322.

[69] Brown, O. W. and Hugenholtz, C. H., 2013. Quantifying the effects of terrestrial laser scanner settings and survey configuration on land surface roughness measurement. *Geosphere*, 9(2):367–377.

[70] Pierzchała, M., Talbot, B., and Astrup, R., 2014. Estimating soil displacement from timber extraction trails in steep terrain: application of an unmanned aircraft for 3D modelling. *Forests*, 5(6):1212–1223.

[71] Borz, S. A. and Proto, A. R., 2022. Application and accuracy of smart technologies for measurements of roundwood: Evaluation of time consumption and efficiency. *Computers and Electronics in Agriculture*, 197:106990.

[72] Borz, S. A. and Proto, A. R., 2025. Low-cost phone-based LiDAR scanning technology provides sub-centimeter accuracy when measuring the main dimensions of motor-manual tree felling cuts. *Ecological Informatics*, 85:102999.

[73] Nevalainen, P., Salmivaara, A., Ala-Ilomäki, J., Launiainen, S., Hiedanpää, J., Finér, L., Pahikkala, T., and Heikkonen, J., 2017. Estimating the rut depth by UAV photogrammetry. *Remote Sensing*, 9(12):1279.

[74] Talbot, B., Rahlf, J., and Astrup, R., 2018. An operational UAV-based approach for stand-level assessment of soil disturbance after forest harvesting. *Scandinavian Journal of Forest Research*, 33(4):387–396.

[75] Büyüksakallı, H., Taş, İ., and Akay, A. E., 2025. Assessing Rut Depth Formation in a Skidding Operation Using UAV Photogrammetry. *European Journal of Forest Engineering*, 11(2):184–191.

[76] Stoffels, J., Hill, J., Sachtleber, T., Mader, S., Buddenbaum, H., Stern, O., Langshausen, J., Dietz, J., and Ontrup, G., 2015. Satellite-Based Derivation of High-Resolution Forest Information Layers for Operational Forest Management. *Forests*, 6(6):1982–2013. ISSN 1999-4907.

[77] Morin, D., Planells, M., Guyon, D., Villard, L., Mermoz, S., Bouvet, A., Thevenon, H., Dejoux, J.-F., Le Toan, T., and Dedieu, G., 2019. Estimation and mapping of forest structure parameters from open access satellite images: Development of a generic method with a study case on coniferous plantation. *Remote Sensing*, 11(11):1275.

[78] Zhichkina, L., Nosov, V., Zhichkin, K., Aydinov, H., Zhenzhebir, V., and Kudryavtsev, V., 2020. Satellite monitoring systems in forestry. In *Journal of Physics: Conference Series*, volume 1515, page 032043. IOP Publishing.

[79] Brožová, N., Krejčí, T., Najmanová, P., Trochta, J., Kučera, M., and Adam, D., 2021. Multiscale analysis of surface roughness for the improvement of natural hazard modelling. *Natural Hazards and Earth System Sciences Discussions*, pages 1–37.

[80] Mohtashami, S., Eliasson, L., Jansson, G., and Sonesson, J., 2017. Influence of soil type, cartographic depth-to-water, road reinforcement and traffic intensity on rut formation in logging operations: a survey study in Sweden. *Silva Fennica*, 51(5).

[81] Grube, G., Grigolato, S., Ala-Ilomäki, J., Routa, J., Lindeman, H., Astrup, R., and Talbot, B., 2025. Modelling machine-induced soil deformation in forest soils using stump proximity and machine learning. *Biosystems Engineering*, 258:104255.

[82] Marra, E., Wictorsson, R., Bohlin, J., Marchi, E., and Nordfjell, T., 2021. Remote measuring of the depth of wheel ruts in forest terrain using a drone. *International Journal of Forest Engineering*, 32(3):224–234.

[83] Grube, G., Talbot, B., and Grigolato, S., 2025. Automated terrain roughness assessment using remotely sensed data. *International Journal of Forest Engineering*, pages 1–16.

[84] Trevisani, S., Teza, G., and Guth, P. L., 2023. Hacking the topographic ruggedness index. *Geomorphology*, 439:108838.

[85] Stefanoni, W., Tocci, D., Latterini, F., Venanzi, R., Gaglioppa, P., Pari, L., and Picchio, R., 2023. A Preliminary Validation and Assessment of a GIS Approach Related to Precision Forest Harvesting in Central Italy. *Forests*, 14(1). ISSN 1999-4907.

[86] Guerra, F., Marzini, S., Sforza, F., Wagner, T., Marinello, F., Grigolato, S., et al., 2024. Exploring the reliability of CAN-bus data in assessing forwarder rolling resistance under real working conditions. *iForest*, 17(6):360–369.

[87] Grassi, G., House, J., Dentener, F., Federici, S., den Elzen, M., and Penman, J., 2017. The key role of forests in meeting climate targets requires science for credible mitigation. *Nature Climate Change*, 7(3):220–226.

[88] Favero, A., Daigneault, A., and Sohngen, B., 2020. Forests: Carbon sequestration, biomass energy, or both? *Science Advances*, 6(13):eaay6792. doi: 10.1126/sciadv.aay6792.

[89] Blattert, C., Mönkkönen, M., Burgas, D., Di Fulvio, F., Toraño Caicoya, A., Vergarechea, M., Klein, J., Hartikainen, M., Antón-Fernández, C., Astrup, R., et al., 2023. Climate targets in European timber-producing countries conflict with goals on forest ecosystem services and biodiversity. *Communications Earth & Environment*, 4(1):119.

[90] Lerink, B. J. W., Schelhaas, M.-J., Schreiber, R., Aurenhammer, P., Kies, U., Vuillermoz, M., Ruch, P., Pupin, C., Kitching, A., Kerr, G., et al., 2023. How much wood can we expect from European forests in the near future? *Forestry: An International Journal of Forest Research*, 96(4):434–447.

[91] Piirainen, S., Finér, L., Andersson, E., Armolaitis, K., Belova, O., Čiuldienė, D., Futter, M., Gil, W., Glazko, Z., Hiltunen, T., Högbom, L., Janek, M., Joensuu, S., Jägrud, L., Libiete, Z., Lode, E., Löfgren, S., Pierzgalski, E., Sikström, U., Zarins, J., and Thorell, D., 2017. Forest drainage and water protection in the Baltic Sea Region countries—current knowledge, methods and areas for development. Technical report, Interreg Baltic Sea Region, European Regional Development Fund. URL https://www.skogsstyrelsen.se/globalassets/projektwebbplatser/wambaf/drainage/reviews/forest-drainage_short_document_imposed_21032017.pdf.

[92] Loisel, J. and Gallego-Sala, A., 2022. Ecological resilience of restored peatlands to climate change. *Communications Earth & Environment*, 3(1):208.

[93] Turunen, J. and Valpola, S., 2020. The influence of anthropogenic land use on Finnish peatland area and carbon stores 1950–2015. *Mires and Peat*, 26:27 pp.

[94] Paavilainen, E. and Päivänen, J., 1995. Utilization of peatlands. In *Peatland forestry: ecology and principles*, pages 15–29. Springer, Berlin, Heidelberg. ISBN 978-3-642-08198-9. doi: 10.1007/978-3-662-03125-4_2.

[95] Peltola, A., 2014. *Metsätilastollinen vuosikirja 2014: Finnish statistical yearbook of forestry 2014*. Metsäntutkimuslaitos (Metla), Vantaa, Finland. ISBN 978-951-40-2506-8. URL https://www.metla.fi/julkaisut/metsatilastollinenvsk/2014/index.htm. Retrieved 2024-04-03.

[96] Laine, J., Laiho, R., Minkkinen, K., and Vasander, H., 2006. Forestry and boreal peatlands. In *Boreal peatland ecosystems*, pages 331–357. Springer.

[97] Similä, M., Aapala, K., and Penttinen, J., editors, 2014. *Ecological restoration in drained peatlands: best practices from Finland*. Metsähallitus, Natural Heritage Services, Vantaa, Finland.

[98] Sallinen, A., Tuominen, S., Kumpula, T., and Tahvanainen, T., 2019. Undrained peatland areas disturbed by surrounding drainage: a large scale GIS analysis in Finland with a special focus on aapa mires. *Mires and Peat*, 24:38.

[99] Zoltai, S. C., Martikainen, P. J., et al., 1996. Estimated extent of forested peatlands and their role in the global carbon cycle. In *NATO ASI Series I: Global Environmental Change*, volume 40, pages 47–58. Springer.

[100] Lepilin, D., Laurén, A., Uusitalo, J., Fritze, H., Laiho, R., Kimura, B., and Tuittila, E.-S., 2022. Response of vegetation and soil biological properties to soil deformation in logging trails of drained boreal peatland forests. *Canadian Journal of Forest Research*, 52(4):511–526.

[101] Ala-Ilomäki, J., Högnäs, T., Lamminen, S., and Sirén, M., 2011. Equipping a conventional wheeled forwarder for peatland operations. *International Journal of Forest Engineering*, 22(1):7–13.

[102] Uusitalo, J. and Ala-Ilomäki, J., 2013. The significance of above-ground biomass, moisture content and mechanical properties of peat layer on the bearing capacity of ditched pine bogs. *Silva Fennica*, 47(3).

[103] Uusitalo, J., Salomäki, M., and Ala-Ilomäki, J., 2015. The effect of wider logging trails on rut formations in the harvesting of peatland forests. *Croatian Journal of Forest Engineering*, 36(1):125–130.

[104] Solgi, A., Naghdi, R., Labelle, E. R., and Zenner, E. K., 2018. The effects of using soil protective mats of varying compositions and amounts on the intensity of soil disturbances caused by machine traffic. *International Journal of Forest Engineering*, 29(3): 199–207.

[105] Ala-Ilomäki, J., Lindeman, H., Mola-Yudego, B., Prinz, R., Väätäinen, K., Talbot, B., and Routa, J., 2021. The effect of bogie track and forwarder design on rut formation in a peatland. *International Journal of Forest Engineering*, 32(sup1):12–19.

[106] Neri, F., Spinelli, R., and Lyons, J., 2007. Ground pressure forwarder trials: assess benefits in reducing wheel rutting. In *Austro2007/FORMEC'07: Meeting the Needs of Tomorrows' Forests – New Developments in Forest Engineering, Vienna and Heiligenkreuz, Austria, October 7–11, 2007*, pages 1–10. URL https://www.formec.org/images/proceedings/2007/session_5_pdf/5_2_paper_neri_spinelli_lyons_austro_formec_2007.pdf.

[107] Brennensthul, M., Czarnecki, J., and Białczyk, W., 2024. Assessment of tractor tires used in forest conditions in terms of traction performance and impact on ground. *Croatian Journal of Forest Engineering*, 45(1):97–114.

[108] Breinig, L., Hinte, B., Schönauer, M., Hoffmann, S., Brokmeier, H., and Jaeger, D., 2025.

Traction assistance of a forwarder in flat terrain: effects on wheel slip and soil disturbance. *Croatian Journal of Forest Engineering*, 46(1):1–17.

[109] Bhatnagar, S., Puliti, S., Talbot, B., Heppelmann, J. B., Breidenbach, J., and Astrup, R., 2022. Mapping wheel-ruts from timber harvesting operations using deep learning techniques in drone imagery. *Forestry*, 95(5):698–710.

[110] Coutts, M. P., 1983. Root architecture and tree stability. In Atkinson, D., Bhat, K. K. S., Coutts, M. P., Mason, P. A., and Read, D. J., editors, *Tree root systems and their mycorrhizas*, pages 171–188. Springer Netherlands, Dordrecht. ISBN 978-94-009-6833-2. doi: 10.1007/978-94-009-6833-2_18.

[111] Wästerlund, I., 1989. Strength components in the forest floor restricting maximum tolerable machine forces. *Journal of Terramechanics*, 26(2):177–182.

[112] Schwarz, M., Giadrossich, F., and Cohen, D., 2013. Modeling root reinforcement using a root-failure Weibull survival function. *Hydrology and Earth System Sciences*, 17 (11):4367–4377.

[113] Ala-Ilomäki, J., 2013. Spiked shear vane—a new tool for measuring peatland top layer strength. *Mires and Peat*, 64(2–3):113–118.

[114] Schwarz, M., Lehmann, P., and Or, D., 2010. Quantifying lateral root reinforcement in steep slopes—from a bundle of roots to tree stands. *Earth Surface Processes and Landforms*, 35(3):354–367.

[115] Kalliokoski, T., 2011. Root system traits of Norway spruce, Scots pine, and silver birch in mixed boreal forests: an analysis of root architecture, morphology, and anatomy. *Dissertationes Forestales*, 121:67.

[116] Stokes, A. and Mattheck, C., 1996. Variation of wood strength in tree roots. *Journal of Experimental Botany*, 47(5):693–699.

[117] Waldron, L. J., 1977. The shear resistance of root-permeated homogeneous and stratified soil. *Soil Science Society of America Journal*, 41(5):843–849.

[118] Wu, T. H., McKinnell, W. P., and Swanston, D. N., 1979. Strength of tree roots and landslides on Prince of Wales Island, Alaska. *Canadian Geotechnical Journal*, 16(1): 19–33.

[119] Genet, M., Stokes, A., Salin, F., Mickovski, S. B., Fourcaud, T., Dumail, J.-F., and van Beek, R., 2005. The influence of cellulose content on tensile strength in tree roots. *Plant and Soil*, 278:1–9.

[120] Piskunov, M., 2023. Influence of stump-root system of trees on rut formation during forwarder operation on peat soils. *Croatian Journal of Forest Engineering*, 44(2): 217–231.

[121] Finnish Meteorological Institute, 2018. Climate in Finland. Retrieved from https://en.ilmatieteenlaitos.fi/climate. Accessed 2024-07-11.

[122] Agisoft LLC, 2021. *Agisoft Metashape user manual: Professional edition, version 1.7*. Retrieved from https://www.agisoft.com/pdf/metashape-pro_1_7_en.pdf.

[123] CloudCompare, 2024. CloudCompare (version 2.13.2) [computer software]. Retrieved from http://www.cloudcompare.org/.

[124] Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., and Yan, G., 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, 8 (6):501.

[125] Miljøministeriet, 2005. Miljøkrav til skovmaskiner på skov- og naturstyrelsens arealer [Environmental requirements for forestry machines on forest and nature agency lands]. Retrieved 2024-12-18. https://naturstyrelsen.dk/media/ysdnrry5/bilag7miljkrav_skovmaskiner2005.pdf.

[126] Masi, E. B., Segoni, S., and Tofani, V., 2021. Root reinforcement in slope stability models: a review. *Geosciences*, 11(5):212.

[127] Schwarz, M., Phillips, C., Marden, M., McIvor, I. R., Douglas, G. B., and Watson, A., 2016. Modelling of root reinforcement and erosion control by 'Veronese' poplar on pastoral hill country in New Zealand. *New Zealand Journal of Forestry Science*, 46: 1–17.

[128] Glass, G. V. and Hakstian, A. R., 1969. Measures of association in comparative experiments: their development and interpretation. *American Educational Research Journal*, 6(3):403–414.

[129] Carroll, R. M. and Nordholm, L. A., 1975. Sampling characteristics of Kelley's ε and Hays' ω. *Educational and Psychological Measurement*, 35(3):541–554.

[130] Breiman, L., 2001. Random forests. *Machine Learning*, 45:5–32.

[131] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[132] Roberts, R., Inzerillo, L., and Di Mino, G., 2020. Using UAV based 3D modelling to provide smart monitoring of road pavement conditions. *Information*, 11(12):568.

[133] van Beek, L. P. H., Wint, J., Cammeraat, L. H., and Edwards, J. P., 2007. Observation and simulation of root reinforcement on abandoned Mediterranean slopes. In *Eco- and ground bio-engineering: the use of vegetation to improve slope stability. Proceedings of the First International Conference on Eco-Engineering, 13–17 September 2004*, pages 91–109. Springer.

[134] Alani, A. M., Ciampoli, L., Bianchini, L., Lantini, L., Tosti, F., and Benedetto, A., 2018. Mapping the root system of matured trees using ground penetrating radar. In *2018 17th International Conference on Ground Penetrating Radar (GPR)*, pages 1–6. IEEE.

[135] Parajuli, M., Gautam, A. P., and Saatchi, S. S., 2022. Logging operations and soil compaction. *Environmental Research Letters*. doi: 10.1088/1748-9326/ac5bb1.

[136] Nordfjell, T., Björheden, R., Thor, M., and Wästerlund, I., 2010. Changes in technical performance, mechanical availability and prices of machines used in forest operations in Sweden from 1985 to 2010. *Scandinavian Journal of Forest Research*, 25(4): 382–389. doi: 10.1080/02827581.2010.498385.

[137] Uusitalo, J., Salomäki, M., and Ala-Ilomäki, J., 2015. Variation of the factors affecting soil bearing capacity of ditched pine bogs in southern Finland. *Scandinavian Journal of Forest Research*, 30(5):429–439. doi: 10.1080/02827581.2015.1012110.

[138] Panagiotidis, D., Abdollahnejad, A., and Slavík, M., 2022. 3D point cloud fusion from UAV and TLS to assess temperate managed forest structures. *International Journal of Applied Earth Observation and Geoinformation*, 112:102917.

[139] Cavalli, M. and Marchi, L., 2008. Characterisation of the surface morphology of an alpine alluvial fan using airborne LiDAR. *Natural Hazards and Earth System Sciences*, 8(2):323–333.

[140] Pollyea, R. M. and Fairley, J. P., 2011. Estimating surface roughness of terrestrial laser scan data using orthogonal distance regression. *Geology*, 39(7):623–626.

[141] Różycka, M., Migoń, P., and Michniewicz, A., 2017. Topographic Wetness Index and Terrain Ruggedness Index in geomorphic characterisation of landslide terrains, on examples from the Sudetes, SW Poland. *Zeitschrift für Geomorphologie Supplementary Issues*, 61(2):61–80.

[142] Lindsay, J. B., Newman, D. R., and Francioni, A., 2019. Scale-optimized surface roughness for topographic analysis. *Geosciences*, 9(7):322.

[143] Sappington, J. M., Longshore, K. M., and Thompson, D. B., 2007. Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. *Journal of Wildlife Management*, 71(5):1419–1426.

[144] Riley, S. J., DeGloria, S. D., and Elliot, R., 1999. Index that quantifies topographic

heterogeneity. *intermountain Journal of sciences*, 5(1-4):23–27.

[145] Hansson, L., Šimůnek, J., Ring, E., Bishop, K., and Gärdenäs, A. I., 2019. Soil compaction effects on root-zone hydrology and vegetation in boreal forest clearcuts. *Soil Science Society of America Journal*, 83:S105–S115.

[146] Svenson, G. and Fjeld, D., 2016. The impact of road geometry and surface roughness on fuel consumption of logging trucks. *Scandinavian Journal of Forest Research*, 31 (5):526–536.

[147] Mohieddinne, H., Brasseur, B., Gallet-Moron, E., Lenoir, J., Spicher, F., Kobaissi, A., and Horen, H., 2023. Assessment of soil compaction and rutting in managed forests through an airborne LiDAR technique. *Land Degradation & Development*, 34(5): 1558–1569.

[148] Doumit, J. A., 2018. Evaluation of multiscale terrain roughness based on UAV datasets: a case of a Lebanese region. In *InterCarto/InterGIS International Conference*, pages 241–249, Petrozavodsk (Russia); Bonn (Germany); Anchorage (USA).

[149] Wallin, E., Wiberg, V., Vesterlund, F., Holmgren, J., Persson, H. J., and Servin, M., 2022. Learning multiobjective rough terrain traversability. *Journal of Terramechanics*, 102: 17–26.

[150] Milenković, M., Pfeifer, N., and Glira, P., 2015. Applying terrestrial laser scanning for soil surface roughness assessment. *Remote Sensing*, 7(2):2007–2045.

[151] Jones, L. and Hobbs, P., 2021. The application of terrestrial LiDAR for geohazard mapping, monitoring, and modelling in the British Geological Survey. *Remote Sensing*, 13(3):395.

[152] Suvinen, A., Tokola, T., and Saarilahti, M., 2009. Terrain trafficability prediction with GIS analysis. *Forest Science*, 55(5):433–442.

[153] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

[154] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., et al., 2023. Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4015–4026.

[155] Nanni, L., Lumini, A., Brahimi, F., Meteier, Q., Sadeghi-Seresht, S., and Logeswaran, R., 2023. Improving existing segmentators performance with zero-shot segmentators. *Entropy*, 25(11):1502.

[156] Kruger, F. J., 1984. Effects of fire on vegetation structure and dynamics. In *Ecological effects of fire in South African ecosystems*, pages 219–243. Springer.

[157] Edwards, D., 1984. Fire regimes in the biomes of South Africa. In *Ecological effects of fire in South African ecosystems*, pages 19–37. Springer.

[158] García-Quijano, J. F., Peters, J., Cockx, L., Van Wyk, G., Rosanov, A., Deckmyn, G., Ceulemans, R., Ward, S. M., Holden, N. M., Van Orshoven, J., et al., 2007. Carbon sequestration and environmental effects of afforestation with Pinus radiata D. Don in the Western Cape, South Africa. *Climatic Change*, 83:323–355.

[159] Rapidlasso GmbH, 2025. *LAStools software including lasground*. URL `https://rapidlasso.com/lastools`. Professional license, version v2.0.3 accessed October 2025.

[160] McGaughey, R. J., September 2023. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Seattle, WA. FUSION Version 4.50.

[161] Bhadoria, P., Agrawal, S., and Pandey, R., 2020. Image segmentation techniques for remote sensing satellite images. In *IOP Conference Series: Materials Science and*

Engineering, volume 993, page 012050. IOP Publishing.

[162] Wada, K., 2021. Labelme: Image Polygonal Annotation with Python. URL https://github.com/wkentaro/labelme. License: GPL-3.

[163] Hobson, R. D., 1967. *FORTRAN IV programs to determine surface roughness in topography for the CDC 3400 computer*. Number 14 in Computer Contribution. State Geological Survey, University of Kansas, Lawrence.

[164] McNicol, I. M., Mitchard, E. T., Aquino, C., Burt, A., Carstairs, H., Dassi, C., Modinga Dikongo, A., and Disney, M. I., 2021. To what extent can UAV photogrammetry replicate UAV LiDAR to determine forest structure? A test in two contrasting tropical forests. *Journal of Geophysical Research: Biogeosciences*, 126(12): e2021JG006586.

[165] Chen, Z., Gao, B., and Devereux, B., 2017. State-of-the-art: DTM generation using airborne LIDAR data. *Sensors*, 17(1):150.

[166] Hell, M., Spengler, D., Frick, A., Wolf, B., and Schulz, J., 2022. Classification of tree species and standing dead trees with lidar point clouds using two deep neural networks: PointCNN and 3DmFV-Net. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90(2):103–121. doi: 10.1007/s41064-022-00200-4.

[167] Singh, A. and Gaurav, K., 2023. Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images. *Scientific Reports*, 13(1):2251.

[168] Pandur, Z., Šušnjar, M., Bačić, M., Đuka, A., Lepoglavec, K., and Nevečerel, H., 2019. Fuel consumption comparison of two forwarders in lowland forests of pedunculate oak. iForest 12: 125–131.

[169] Smidt, M. F., Thompson, J., McDonald, T., and Ma, Y., 2023. Terra-mechanical and statistical analyses of four-wheel grapple skidder performance data. *Research Paper SRS–66. Asheville, NC: US Department of Agriculture, Forest Service, Southern Research Station*, 66:1–16.

[170] Orlof, J., Ozimek, P., Łabędź, P., Widłak, A., and Ozimek, A., 2024. Generating viewsheds based on the Digital Surface Model (DSM) and point cloud. *PLOS ONE*, 19(12): e0312146.

[171] Ryding, J., October 2016. *Assessing New Methods for Measuring Forest Understorey Vegetation Using Terrestrial Laser Scanning*. Phd thesis, University of Nottingham.

[172] Qi, C. R., Yi, L., Su, H., and Guibas, L. J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In *Advances in Neural Information Processing Systems*. URL https://arxiv.org/abs/1706.02413.

[173] Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., and Markham, A., 2020. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11108–11117.

[174] Lu, H., Li, B., Yang, G., Fan, G., Wang, H., Pang, Y., Wang, Z., Lian, Y., Xu, H., and Huang, H., 2025. Towards a point cloud understanding framework for forest scene semantic segmentation across forest types and sensor platforms. *Remote Sensing of Environment*, 318:114591.

[175] Laino, D., Cabo, C., Ordóñez, C., Bolaños, R., Janvier, R., Giulioni, F., Herrmann, M., Hudak, A., Parsons, R., and Santin, C., 2025. SegmentedForests: a labelled dataset of terrestrial LiDAR point clouds for semantic segmentation of forests. *Forestry: An International Journal of Forest Research*, page cpaf062.

[176] Puliti, S., Pearse, G., Surovỳ, P., Wallace, L., Hollaus, M., Wielgosz, M., and Astrup, R., 2023. For-instance: a uav laser scanning benchmark dataset for semantic and instance segmentation of individual trees. *arXiv preprint arXiv:2309.01279*.

[177] Henrich, J., van Delden, J., Seidel, D., Kneib, T., and Ecker, A. S., 2024. TreeLearn: A deep learning method for segmenting individual trees from ground-based LiDAR forest point clouds. *Ecological Informatics*, 84:102888.

[178] RIEGL Laser Measurement Systems GmbH, 2020. RiSCAN PRO. URL https://www.riegl.com/products/software-packages/riscan-pro/. Software used for TLS data processing.

[179] The MathWorks Inc., 2025. MATLAB version: R2025b. URL https://www.mathworks.com.

[180] Vicari, M. B., Disney, M., Wilkes, P., Burt, A., Calders, K., and Woodgate, W., 2019. Leaf and wood classification framework for terrestrial LiDAR point clouds. *Methods in Ecology and Evolution*, 10(5):680–694. ISSN 2041-210X.

[181] Nurunnabi, A., Teferle, F., Laefer, D. F., Chen, M., and Ali, M. M., 2024. Development of a precise tree structure from LiDAR point clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48:301–308.

[182] Amakhchan, W., Tarsha Kurdi, F., Gharineiat, Z., Boulaassal, H., and El Kharki, O., 2023. Classification of Forest LiDAR Data Using Deep Learning Pipeline Algorithm and Geometric Feature Analysis. *International Journal of Environmental Sciences Natural Resources*, 32(3):556340.

[183] Kulicki, M., Cabo, C., Trzciński, T., Będkowski, J., and Stereńczak, K., 2025. Artificial Intelligence and Terrestrial Point Clouds for Forest Monitoring. *Current Forestry Reports*, 11(5):5. doi: 10.1007/s40725-024-00234-4. Accepted 26 September 2024.

[184] Ruoppa, L., Hietala, T., Seppänen, V., Taher, J., Hakala, T., Yu, X., Kukko, A., Kaartinen, H., and Hyyppä, J., 2025. Benchmarking individual tree segmentation using multispectral airborne laser scanning data: the FGI-EMIT dataset. *arXiv preprint arXiv:2511.00653*. URL https://arxiv.org/abs/2511.00653.

[185] Zhang, F., Chancia, R., Clapp, J., Hassanzadeh, A., Dera, D., MacKenzie, R., and van Aardt, J., 2025. Through the Perspective of LiDAR: A Feature-Enriched and Uncertainty-Aware Annotation Pipeline for Terrestrial Point Cloud Segmentation. *arXiv preprint arXiv:2510.06582*.

[186] Kim, Y., Yoon, D., Lee, Y., Fatima, U., Kim, H. K., Lee, S., Park, S., Park, J. H., Kang, S., and Jeon, M., 2025. Unlocking Robust Semantic Segmentation Performance via Label-only Elastic Deformations against Implicit Label Noise. *arXiv preprint arXiv:2508.10383*. URL https://arxiv.org/abs/2508.10383v2.

[187] Price, M. F., Gratzer, G., Alemayehu Duguma, L., Kohler, T., Maselli, D., and Romeo, R., editors, 2011. *Mountain Forests in a Changing World: Realizing Values, Addressing Challenges*. Food and Agriculture Organization of the United Nations (FAO) and Swiss Agency for Development and Cooperation (SDC), Rome. ISBN 978-92-5-107076-5. International Year of Forests publication.

[188] Egan, P. A. and Price, M. F., editors, 2017. *Mountain ecosystem services and climate change: a global overview of potential threats and strategies for adaptation*. UNESCO, Paris. ISBN 978-92-3-100225-0. URL https://unesdoc.unesco.org/ark:/48223/pf0000248768.

[189] Donald, K., Boswell, B., Amishev, D., and Hunt, J., June 2018. Winch-Assist Harvester: Best Practice Manual. Special Publication SP-533, FPInnovations. URL /mnt/data/WinchAssistHarvesterBMP_compressed.pdf. Print ISBN 978-0-86488-578-4; digital ISBN 978-0-86488-579-1.

[190] Warren, S. D., Hohmann, M. G., Auerswald, K., and Mitasova, H., 2004. An evaluation of methods to determine slope using digital elevation data. *Catena*, 58(3):215–233. ISSN 0341-8162.

[191] Lakshmi, S. E. and Yarrakula, K., 2018. Review and critical analysis on digital ele-

vation models. *Geofizika*, 35(2):129–157. doi: 10.15233/gfz.2018.35.7. Review paper.

[192] Hartley, R. J. L., Jayathunga, S., Elleouet, J. S., Steer, B. S. C., and Watt, M. S., 2025. UAV-enabled evaluation of forestry plantations: A comprehensive assessment of laser scanning and photogrammetric approaches. *Science of Remote Sensing*, 12:100245. ISSN 2666-0172.

[193] Shao, J., Zhang, W., Luo, L., Cai, S., and Jiang, H., 2020. SLAM-Based Backpack Laser Scanning for Forest Plot Mapping. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-2-2020:267–271.

[194] Balenović, I., Liang, X., Jurjević, L., Hyyppä, J., Seletković, A., and Kukko, A., 2021. Hand-held personal laser scanning–current status and perspectives for forest inventory application. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 42(1):165–183.

[195] Wilson, M. F., O'Connell, B., Brown, C., Guinan, J. C., and Grehan, A. J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30(1-2):3–35.

[196] Spadini, G., October 2022. Terna – Monitoraggio infrastruttura: Linee aeree. Presentation, Asset Management.

[197] Terna Driving Energy S.p.A., 2025. Annual Report 2024. Approved by the Board of Directors on 25 March 2025.

[198] American Society for Photogrammetry and Remote Sensing (ASPRS), 2019. *LAS Specification 1.4 - R14*. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, version 1.4 - r14 edition. URL https://github.com/ASPRSorg/LAS. Version approved November 2011. Revision date: 26 March 2019.

[199] DJI Terra, 2024. *DJI Terra (Version 4.5.0) [Software]*. URL https://enterprise.dji.com/it/dji-terra. Professional license, accessed July 2025.

[200] Stonex SRL., 2022. *X120GO SLAM Laser Scanner User Guide*. Stonex Srl, Paderno Dugnano (MI), Italy. URL https://www.stonex.it. Version 1.0, Rev 1. Accessed August 2025.

[201] Moran, P. A. P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37 (1/2):17–23. doi: 10.1093/biomet/37.1-2.17.

[202] NASA Earthdata, 2025. Digital Elevation/Terrain Model (DEM). URL https://www.earthdata.nasa.gov/topics/land-surface/digital-elevation-terrain-model-dem. Accessed: 2025-12-10.

[203] Ågren, A. M., Lidberg, W., Strömgren, M., Ogilvie, J., and Arp, P. A., 2014. Evaluating digital terrain indices for soil wetness mapping – a Swedish case study. *Hydrology and Earth System Sciences*, 18(9):3623–3634. doi: 10.5194/hess-18-3623-2014.

[204] Waga, K., Malinen, J., and Tokola, T., 2020. A Topographic Wetness Index for Forest Road Quality Assessment: An Application in the Lakeland Region of Finland. *Forests*, 11(11):1165. doi: 10.3390/f11111165.

[205] Horn, B. K. P., 1981. Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1):14–47. doi: 10.1109/PROC.1981.11918. URL https://doi.org/10.1109/PROC.1981.11918.

[206] Böhner, J. and Selige, T., 2006. Spatial prediction of soil attributes using terrain analysis and climate regionalisation. *Gottinger Geographische Abhandlungen*, 115: 13–28.

[207] Moore, I. D., Grayson, R., and Ladson, A., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological processes*, 5(1):3–30.

[208] Flood, M., May 2004. ASPRS Guidelines: Vertical Accuracy Reporting for Lidar Data,

Version 1.0. Technical report, American Society for Photogrammetry and Remote Sensing (ASPRS) Lidar Committee (PAD), Bethesda, MD. Released May 24, 2004. Drafted May 15, 2004.

[209] Visser, R. and Stampfer, K., 2015. Expanding ground-based harvesting onto steep terrain: A review. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 36(2):321–331.

[210] Salach, A., Bakuła, K., Pilarska, M., Ostrowski, W., Górski, K., and Kurczyński, Z., 2018. Accuracy assessment of point clouds from LiDAR and dense image matching acquired using the UAV platform for DTM creation. *ISPRS International Journal of Geo-Information*, 7(9):342.

[211] Sappington, J. M., Longshore, K. M., and Thompson, D. B., 2007. Quantifying landscape ruggedness for animal habitat analysis. *Journal of Wildlife Management*.

[212] Garland, J., Belart, F., Crawford, R., Chung, W., Cushing, T., Fitzgerald, S., Green, P., Kincl, L., Leshchinsky, B., Morrissette, B., et al., 2019. Safety in steep slope logging operations. *Journal of agromedicine*, 24(2):138–145.

[213] Meng, X., Currit, N., and Zhao, K., 2010. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sensing*, 2(3):833–860.

[214] Silva, C. A., Klauberg, C., Hentz, Â. M. K., Corte, A. P. D., Ribeiro, U., and Liesenberg, V., 2018. Comparing the performance of ground filtering algorithms for terrain modeling in a forest environment using airborne LiDAR data. *Floresta e Ambiente*, 25(2): e20160150.

[215] Klápště, P., Fogl, M., Barták, V., Gdulová, K., Urban, R., and Moudrỳ, V., 2020. Sensitivity analysis of parameters and contrasting performance of ground filtering algorithms with UAV photogrammetry-based and LiDAR point clouds. *International Journal of Digital Earth*, 13(12):1672–1694.

[216] Stereńczak, K., Ciesielski, M., Balazy, R., and Zawiła-Niedźwiecki, T., 2016. Comparison of various algorithms for DTM interpolation from LIDAR data in dense mountain forests. *European Journal of Remote Sensing*, 49(1):599–621.

[217] Habib, M., 2021. Evaluation of DEM interpolation techniques for characterizing terrain roughness. *Catena*, 198:105072.

[218] Dainelli, R., Toscano, P., Di Gennaro, S. F., and Matese, A., 2021. Recent Advances in Unmanned Aerial Vehicle Forest Remote Sensing—A Systematic Review. Part I: A General Framework. *Forests*, 12(3). ISSN 1999-4907. URL https://www.mdpi.com/1999-4907/12/3/327.

[219] Fraser, B. T., Bunyon, C. L., Reny, S., Lopez, I. S., and Congalton, R. G., 2022. Analysis of unmanned aerial system (UAS) sensor data for natural resource applications: A review. *Geographies*, 2(2):303–340.

[220] Pavelka, K., Matoušková, E., and Pavelka Jr, K., 2023. Remarks on Geomatics Measurement Methods Focused on Forestry Inventory. *Sensors*, 23(17):7376.

[221] Puliti, S., Ørka, H. O., Gobakken, T., and Næsset, E., 2015. Inventory of Small Forest Areas Using an Unmanned Aerial System. *Remote Sensing*, 7(8):9632–9654. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/7/8/9632.

[222] Nekhin, S., Rubenok, A., and Kovrov, A., 2024. Research on the Effectiveness of Lidar Survey for Large Mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48:529–534.

[223] Piermattei, L., Karel, W., Wang, D., Wieser, M., Mokroš, M., Surový, P., Koreň, M., Tomaštík, J., Pfeifer, N., and Hollaus, M., 2019. Terrestrial Structure from Motion Photogrammetry for Deriving Forest Inventory Data. *Remote Sensing*, 11(8). ISSN 2072-4292. doi: 10.3390/rs11080950. URL https://www.mdpi.com/2072-4292/11/8/950.

[224] Ma, Y., Wu, H., Wang, L., Huang, B., Ranjan, R., Zomaya, A., and Jie, W., 2015. Remote sensing big data computing: Challenges and opportunities. *Future Generation Computer Systems*, 51:47–60.

[225] NASA Earthdata, 2024. ESDS Data Metrics. https://www.earthdata.nasa.gov/about/data-metrics. Accessed: 2025-11-22.

[226] ReOrbit, 2025. Earth Observation Today: Rapid data from space unlocking new business opportunities. Available online. Accessed: 2025-11-22.

[227] Maurya, N. K., Tripathi, A. K., Chauhan, A., Pandey, P. C., and Lamine, S., 2022. Recent advancement and role of drones in forest monitoring: Research and practices. *Advances in Remote Sensing for Forest Monitoring*, pages 221–254.

[228] Xu, C., Du, X., Fan, X., Giuliani, G., Hu, Z., Wang, W., Liu, J., Wang, T., Yan, Z., Zhu, J., et al., 2022. Cloud-based storage and computing for remote sensing big data: a technical review. *International Journal of Digital Earth*, 15(1):1417–1445.

[229] Roosevelt, C. H., 2014. Mapping site-level microtopography with real-time kinematic global navigation satellite systems (RTK GNSS) and unmanned aerial vehicle photogrammetry (UAVP). *Open Archaeology*, 1(1).

[230] Krisanski, S., Taskhiri, M. S., and Turner, P., 2020. Enhancing Methods for Under-Canopy Unmanned Aircraft System Based Photogrammetry in Complex Forests for Tree Diameter Measurement. *Remote Sensing*, 12(10). ISSN 2072-4292. doi: 10.3390/rs12101652. URL https://www.mdpi.com/2072-4292/12/10/1652.

[231] Borgogno-Mondino, E., De Petris, S., Sarvia, F., Momo, E. J., Sussio, F., and Pari, P., 2022. Adoption of Digital Aerial Photogrammetry in Forest Planning: A Case Study of Canavese Forestry Consortium, NW Italy with Technical and Economic Issues. *Land*, 11(8). ISSN 2073-445X. URL https://www.mdpi.com/2073-445X/11/8/1350.

[232] DJI, 02 2016. DJI Adjusts Pricing for Phantom 3 Professional Globally. https://www.dji.com/newsroom/news/dji-adjusts-price-for-phantom-3-professional. Press release.

[233] DJI, 2025. DJI Mini 4K. https://store.dji.com/lv/product/dji-mini-4k. Product page.

[234] Delinc, J., 2015. Technical Report on Cost Effectiveness of Remote Sensing for Agricultural Statistics in Developing and Emerging Economies. Technical report, Food and Agriculture Organization of the United Nations (FAO). URL https://www.gsars.org. Prepared in the framework of the Global Strategy to Improve Agricultural and Rural Statistics.

[235] Foster, A., Rahimzadeh-Bajgiran, P., Daigneault, A., and Weiskittel, A., 2024. Cost-effectiveness of remote sensing technology for spruce budworm monitoring in Maine, USA. *Forests Monitor*, 1(1):66–98.

[236] Ravio, 2025. The average pay rise in 2025: Europe edition. URL https://ravio.com/blog/average-pay-rise. Accessed: 2025-11-21.

[237] Kvirikadze, I., 2025. Global Outlook: Expected Salary Increases for 2025. https://www.venconresearch.com/inisght/global-outlook-expected-salary-increases-for-2025. Accessed: 2025-11-21.

[238] European Flyers, 2024. Drone/RPAS pilot course - European Flyers. URL https://europeanflyers.com/en/drone-pilot-course-rpa. Drone pilot training and certification cost information, Europe.

[239] The Academy of Aviation, 2025. RPC - Remote Pilot Certificate Licence (Drone Pilot) - The Academy of Aviation, South Africa. URL https://www.theacademyofaviation.co.za/product/rpc-remote-pilot-licence-3-weeks-full-time/.

[240] RocketDNA, 2025. Remote Pilot Certificate (RPC) Course - RocketDNA. URL https://www.rocketdna.com/course/remote-pilot-certificate-course.

[241] Evans, J. S., Hudak, A. T., Faux, R., and Smith, A. M. S., 2009. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sensing*, 1(4):776–794. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/1/4/776.

[242] Moretto, J., 2015. Activity report for the research grant: Innovative tools for land management and planning of small private forest holdings. Technical report within the For.Mo.Vi. project. Preliminary report on UAV and LiDAR surveys in forest areas of Valli del Pasubio and Tortima (VI), Italy.

[243] POLIS Network, 2025. European Commission Shares First Findings on the True Costs of EU Transport. Available online. Accessed: 2025-11-21.

[244] Directory, P. S., 2025. Urban Logistics Greening with Drones. URL https://prism.sustainability-directory.com/scenario/urban-logistics-greening-with-drones/. Accessed: 2025-11-21.

[245] DJI, 2025. Supporto per Matrice 300 RTK - DJI Italia. URL https://www.dji.com/it/support/product/matrice-300. Accessed: 2025-11-21.

[246] Dunham, S., June 2025. *TxDOT Airborne Lidar Specifications and Project Workflow for Design-Grade Mapping Applications*. TxDOT Design Division Remote Sensing Services Section and Gorrondona & Associates, Inc. URL https://www.txdot.gov/content/dam/docs/division/row/survey/airborne-lidar-for-design.pdf. Technical Manual.

[247] Gao, M., Hugenholtz, C. H., Fox, T. A., Kucharczyk, M., Barchyn, T. E., and Nesbit, P. R., 2021. Weather constraints on global drone flyability. *Scientific Reports*, 11 (1):12092.

[248] Nicoll, B. C. and Ray, D., 1996. Adaptive growth of tree root systems in response to wind action and site conditions. *Tree Physiology*, 16(11-12):891–898. doi: 10.1093/treephys/16.11-12.891.

[249] Di Iorio, A., Lasserre, B., Scippa, G. S., and Chiatante, D., 2005. Root System Architecture of *Quercus pubescens* Trees Growing on Different Sloping Conditions. *Annals of Botany*, 95(2):351–361.

[250] Smith, A., Astrup, R., Raumonen, P., Liski, J., Krooks, A., Kaasalainen, S., Åkerblom, M., and Kaasalainen, M., 2014. Tree root system characterization and volume estimation by terrestrial laser scanning and quantitative structure modeling. *Forests*, 5 (12):3274–3294.

[251] Ballard, T. M., 2000. Impacts of forest management on northern forest soils. *Forest Ecology and Management*, 133(1-2):37–42. ISSN 0378-1127.

[252] Kozlowski, T. T., 1999. Soil compaction and growth of woody plants. *Scandinavian Journal of Forest Research*, 14(6):596–619.

[253] Ampoorter, E., De Frenne, P., Hermy, M., and Verheyen, K., 2011. Effects of soil compaction on growth and survival of tree saplings: A meta-analysis. *Basic and Applied Ecology*, 12(5):394–402. doi: 10.1016/j.baae.2011.06.003.

[254] Pyttel, P. L., Fischer, U. F., Suchomel, C., Gärtner, S. M., and Bauhus, J., 2013. The effect of harvesting on stump mortality and re-sprouting in aged oak coppice forests. *Forest Ecology and Management*, 289:18–27.

[255] Spinelli, R., Magagnotti, N., and Schweier, J., 2017. Trends and Perspectives in Coppice Harvesting. *Croatian Journal of Forest Engineering*, 38(2):219–230.

[256] Marinello, F., Proto, A. R., Zimbalatti, G., Pezzuolo, A., Cavalli, R., and Grigolato, S., 2017. Determination of forest road surface roughness by Kinect depth imaging. *Annals of Forest Research*, 60(2):217–226.

[257] Mikleš, M., Helexa, M., and Mikleš, J., 2017. Research on forest terrain roughness as a source of dynamic action on the vehicle. *Journal of Forest Science*, 63(8):363–369.

[258] Dai, W., Yang, X., Na, J., Li, J., Brus, D., Xiong, L., Tang, G., and Huang, X., 2019. Effects of DEM resolution on the accuracy of gully maps in loess hilly areas. *Catena*, 177: 114–125. ISSN 0341-8162.

[259] Li, Y., McNelis, J. J., and Washington-Allen, R. A., 2020. Quantifying Short-Term Erosion and Deposition in an Active Gully Using Terrestrial Laser Scanning: A Case Study From West Tennessee, USA. *Frontiers in Earth Science*, 8:587999.

[260] Wagemann, J., Siemen, S., Seeger, B., and Bendix, J., 2021. Users of open Big Earth data–An analysis of the current state. *Computers & Geosciences*, 157:104916.

[261] Buttlar, W. G., Choi, J., and Lee, J., 2021. Uses and Challenges of Collecting LiDAR Data from a Mobile Mapping System for Roadway Data Collection. Technical Report MPC-21-427, Mountain-Plains Consortium, Upper Great Plains Transportation Institute, North Dakota State University, Fargo, ND. URL https://www.ugpti.org/resources/reports/downloads/mpc21-427.pdf.

[262] Kleist, N. J., Domschke, C. T., Knight, A. C., Nauman, T. W., Duniway, M. C., and Carter, S. K., 2025. Mapping predicted ecological states at landscape scales using remote-sensing data and machine learning. *Ecosphere*, 16(4):e70243.

[263] Strahlendorff, M., Kosmale, M., Ruosteenoja, K., Ziese, M., et al., 2024. Forestry climate adaptation with HarvesterSeasons service: forecasting soil wetness for climate-smart operations. *Frontiers in Remote Sensing*, 5:1360572. doi: 10.3389/frsen.2024.1360572. URL https://www.frontiersin.org/articles/10.3389/frsen.2024.1360572/full.

[264] Arbonaut Ltd, 2026. Forest Operations. https://arbonaut.com/services/forest-operations. Accessed: 2026-01-06.

[265] Swedish Forest Agency, 2024. Digital Open Forest data. https://www.skogsstyrelsen.se/en/digital-open-forest-data/. Accessed: 6 January 2026.

[266] Fassnacht, F. E., White, J. C., Wulder, M. A., and Næsset, E., 2024. Remote sensing in forestry: current challenges, considerations and directions. *Forestry: An International Journal of Forest Research*, 97(1):11–37.

[267] Lundbäck, M., Wallin, E., Häggström, C., Nyström, M., Grönlund, A., Richardson, M., Jönsson, P., Arnvik, W., Hedström, L., Fälldin, A., et al., 2025. FORWARD: Dataset of a forwarder operating in rough terrain. *arXiv preprint arXiv:2511.17318*.

[268] Liang, X., Yao, H., Qi, H., and Wang, X., 2024. Forest in situ observations through a fully automated under-canopy unmanned aerial vehicle. *Geo-Spatial Information Science*, 27(4):983–999.

[269] Mattamala, M., Chebrolu, N., Frey, J., Freißmuth, L., Oh, H., Casseau, B., Hutter, M., and Fallon, M., 2025. Building Forest Inventories with Autonomous Legged Robots–System, Lessons, and Challenges Ahead. *arXiv preprint arXiv:2506.20315*.

[270] Khaksar, W. and Astrup, R., 2021. Multi-Sensor Terrestrial SLAM for Real-Time, Large-Scale, and GNSS-Interrupted Forest Mapping. In *Proceedings of the FORMEC/-Forest Engineering Conference*, Ås, Norway. Faculty of Science and Technology, Norwegian University of Life Sciences (NMBU); Norwegian Institute of Bioeconomy Research (NIBIO).

[271] La Hera, P., Mendoza-Trejo, O., Lindroos, O., Lideskog, H., Lindbäck, T., Latif, S., Li, S., and Karlberg, M., 2024. Exploring the feasibility of autonomous forestry operations: Results from the first experimental unmanned machine. *Journal of Field Robotics*, 41 (4):942–965.

[272] Hellstrom, T. and Ringdahl, O., 2006. Follow the Past: a path-tracking algorithm for

autonomous vehicles. *International Journal of Vehicle Autonomous Systems*, 4(2–4): 216–224.