Enhancement of Digital Elevation Models using Tree-Based Ensemble Machine Learning Algorithms

By:

Chukwuma John Okolie

STUDENT NO: OKLCHU002

Supervisors:

Assoc. Professor Julian Smit (2020 – 2022)

Professor Jennifer Whittal (Administrative)

Dr. Adedayo Adeleke

Submitted to the University of Cape Town in fulfilment of the requirements for the Degree of Doctor of Philosophy in Engineering (Geomatics)
December 2023

Department of Architecture, Planning and Geomatics: Division of Geomatics
Faculty of Engineering and the Built Environment
University of Cape Town

DECLARATION

I, **Chukwuma John Okolie**, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for research either the whole or any portion of the contents in any manner whatsoever.

Signature:

Date: December 2023

DEDICATION

To:

My wonderful parents, the best in the world

James Ugochukwu Okolie *PhD, FCA*

(of blessed memory)

&

Martha Obioma Okolie

ACKNOWLEDGEMENTS

I am grateful to the Almighty God, for His goodness, mercies, and enablement throughout the journey of this research. I thank all my supervisors who have mentored me throughout this journey, Assoc. Professor Julian Smit (who is no longer at the University of Cape Town), Professor Jennifer Whittal and Dr Adedayo Adeleke for their immense support, guidance and devotion to facilitating the successful completion of this research. I also appreciate their excellent and meticulous supervision of my research activity. I am also grateful to Professor Jon Mills who supervised my research conducted at Newcastle University (UK) under the auspices of the Commonwealth split-site PhD programme.

I salute the management of the University of Lagos (UniLag), and the immediate past dean of the Faculty of Engineering at UniLag (Professor Funso Falade) for facilitating my study leave to undertake doctoral research at UCT. Special appreciation to my Professors, senior colleagues and colleagues at the Department of Surveying and Geoinformatics (UniLag) for their support and encouragement: Professor Peter Nwilo (of blessed memory), Dr Dupe Olayinka-Dosunmu, Dr Emmanuel Ayodele, Dr Sunday Alademomi, Dr Isa Hamid-Mosaku, Dr O.E. Abiodun, Mr Michael Orji, Mr Olagoke Daramola, Mr Abiodun Alabi, Mr Tosin Salami, Mrs Funmilola Ogunlade and Mrs Funmilayo Oyatayo. I also appreciate the immediate past Heads of Department (HODs), Assoc Professor Gabriel Omogunloye and Professor Joseph Olusina; and the present HOD, Dr Emeka Epuh. In the course of my research, I benefitted from the knowledge of Mr Ikenna Arungwa, Dr Maria Peppa, Dr Shidong Wang, Mr Tom Komar, Dr Ikechukwu Maduako, Dr Erdogan Taskesen and Dr Hossein Bagheri who were very kind with resources and helped with reviewing my results, and making suggestions for improvement. Dr Jerry Obiefuna has been a good mentor over the years and a supporter of my academic career. I thank my past students at the University of Lagos including the present set who keep asking when I am coming back. A special salutation to Johanson Onyegbula, Caleb Ogbeta, Samuel Akinnusi, Erom Mbu-Ogar, Waliyah Adedokun, Hamed Olanrewaju, Abdullahi Hamzat, Imole Okediji and Andy Egogo-Stanley. I acknowledge the guidance and counsel from Dr Nehemiah Alozie and Professor Samson Adeosun during my programme. I also thank several anonymous journal and conference reviewers whose comments have greatly improved the quality of the research.

At UCT Geomatics Division, Assoc Prof Patroba Odera, Dr Simon Hull, Dr Moreblessings Shoko, Dr Siphiwe Mphuthi, Mr Kaveer Singh, Ms Mignon Wells, Ms Junita Abrahams, Mr

Dirk Matthee and Ms Suleiga Schröder were very helpful, kind and welcoming. Mr Thomas Slingsby (GIS Officer at UCT Libraries) was very kind and forthcoming with ideas for my research. I must also appreciate my fellow postgraduate students at UCT Geomatics Division, Ms Mamphaka Mabesa, Ms Chidinma Ndu, Ms Dianah Abeho, Mr Elly Ngailo, Mr Epie Njume, Mr Abdulganiy Shittu and Mr Alex Yumbu. I benefitted from the useful feedback from my friends, Chima Iheaturu, Abdulwaheed Tella and Lisah Ligono. Special thanks to Prof Abimbola Windapo and Dr Kehinde Babalola for being kind and supportive with useful information as I settled in at UCT; and my noble sisters in the Lord, Kelechi Njoku and Dupe Ojo. Special thanks to Dr Desiree Moodley and Associate Professor Kate le Roux for their tremendous guidance under the 'Writing for Postgraduate Completion' initiative in the Engineering and Built Environment Faculty at UCT.

I am grateful for the financial support provided by the University of Cape Town Postgraduate Funding Office (PGFO) and the Commonwealth Scholarship Commission (CSC-UK). I also appreciate the Deeper Christian Life Ministry community at UCT, and the ministers at the Redeemed Christian Church of God (RCCG) Latter House Parish, Cape Town.

Finally, I appreciate my mom, Mrs Martha Obioma Okolie for her undying love, unwavering devotion, affection, and constant encouragement. I must acknowledge my siblings (Dr Chukwuemeka Okolie, Mr Uchechukwu Okolie, Ms Onyinyechukwu Okolie, Ms Chinyere Okolie and Mr Okechukwu Okolie) for their immense love, encouragement and support. Special thanks to my sister-in-law, Mrs Chinyere Okolie, for being kind and hospitable during my visit to the United Kingdom.

ABSTRACT

Geospatial practitioners and national mapping agencies (NMAs) in Africa are constrained to rely on open-access remote sensing datasets, even as they struggle to meet up with best practices on spatial data infrastructure and topographic map revision. Thus, global digital elevation models (DEMs) have gained worldwide prominence due to their free availability; a prime advantage when compared to prohibitively expensive airborne topographic surveys. However, the accuracies of global DEMs are affected by several anomalies that diminish their quality and compromise their adequacy for applications where precise and accurate terrain information is needed. This research proposes an explainable tree-based ensemble feature-level fusion framework for enhancing satellite DEMs using Cape Town, South Africa as a case study. The enhancement methodology combines elevation and terrain features data alignment (co-registration and resampling) with feature-level fusion (ensemble learning) into a DEM enhancement framework. The training datasets are comprised of eleven predictor variables including elevation, slope, aspect, surface roughness, topographic position index, terrain ruggedness index, terrain surface texture, vector ruggedness measure, percentage bare ground, urban footprints and percentage forest cover as an indicator of the overland forest distribution. The target variable (elevation error) was derived with respect to highly accurate airborne LiDAR. Initially, the qualities of two satellite-derived interferometric DEMs (NASADEM and Copernicus) and two photogrammetric DEMs (ASTER and AW3D) were comparatively examined in a series of qualitative and quantitative tests in five different landscapes spread across Cape Town: urban/industrial, agricultural, mountain, peninsula and grassland/shrubland. Based on their performances, Copernicus and AW3D DEMs were selected for further analysis. The next phase involved a comparative evaluation of ten tree-based ensembles for enhancement of Copernicus DEM over agricultural lands. At two implementation sites, there was a 6 – 13% reduction in the MAE and 15 – 29% reduction in the RMSE, and the corrected Copernicus DEM showed several topographic improvements such as smoothing of rough edges, enhanced stream channel conditioning and diminution of coarse/grainy pixels. Following the comparison, three recent implementations of gradient boosting, the extreme gradient boosting (XGBoost), light boosting machine (LightGBM) and categorical boosting (CatBoost) were selected for the development of a robust DEM enhancement framework. After training and testing, the models were applied for correcting the DEMs at two implementation sites spread across the five landscapes. Going further, a rigorous hyperparameter tuning strategy was implemented for the three models using a principled,

robust and computationally efficient Bayesian optimisation scheme. The optimisations were operationalised with ten steps of random exploration for diversification of the exploration space, and 40 - 50 iterations to increase the likelihood of finding an optimal combination of hyperparameter values. The uniqueness of the optimisation scheme is the very wide diversification of the search space for random exploration. The performance of the models was compared based on default hyperparameters versus Bayesian-optimised hyperparameters. The result is a sequential correction and fusion scheme to increase the vertical accuracy and reduce errors in the final DEMs. The corrections achieved significant and highly competitive accuracy gains of up to 64% RMSE (68% MAE) reduction in Copernicus DEM and up to 78% RMSE (82% MAE) reduction in AW3D DEM. The robustness of the proposed framework was proven in several performance evaluations and comparative assessments. Summarily, it outperformed a globally acclaimed corrected DEM and the authoritative South Africa national DEM, and surpassed the achievable accuracies of several previously proposed strategies, including multiple linear regression. Moreover, three-dimensional terrain analysis and lineament mapping showed the potential of the proposed scheme for enhancing deliverables in topographic and geologic mapping. The proposed approach also incorporates explainability measures to describe the interactions between predictor variables and their influence on the predicted DEM errors. It provides a cost-effective framework and ‘minimal’ computation expense. The innovative DEM enhancement scheme proposed in this research is applicable to other global landscapes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

AdaBoost	Adaptive Boosting
AI	Artificial Intelligence
ALOS	Advanced Land Observing Satellite
ANN	Artificial Neural Network
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AW3D	ALOS World 3D
Bagging	Bootstrap Aggregation
BO	Bayesian Optimisation
CART	Classification and Regression Tree
CatBoost	Categorical Gradient Boosting
CD: NGI	Chief Directorate: National Geospatial Information
COCT	City of Cape Town
DEM	Digital Elevation Model
DGM	Digital Ground Model
DHM	Digital Height Model
DL	Deep Learning
DLR	German Aerospace Centre
DSM	Digital Surface Model
DT	Decision Tree
DTED	Digital Terrain Elevation Data
DTM	Digital Terrain Model
EFB	Exclusive Feature Bundling
EGM	Earth Gravitational Model
ESRI	Environmental Systems Research Institute
ETM	Enhanced Thematic Mapper
FABDEM	Forests and Buildings removed Copernicus DEM

GBDT	Gradient Boosted Decision Tree
GBM	Gradient Boosting Machine
GCP	Ground Control Point
GCS	Geographic Coordinate System
GDEM	Global Digital Elevation Model
GEDI	Global Ecosystem Dynamics Investigation
GIS	Geographic Information System
GLAD	Global Land Analysis and Discovery
GLAS	Geoscience Laser Altimeter System
GNSS	Global Navigation Satellite Systems
GOSS	Gradient-based One-Side Sampling
GPS	Global Positioning System
GUF	Global Urban Footprint
HEM	Height Error Map
IAG	International Association of Geodesy
ICESat	Ice, Cloud and Land Elevation Satellite
ICGEM	International Centre for Global Earth Models
IGFS	International Gravity Field Service
INSAR	Interferometric Synthetic Aperture Radar
JAXA	Japan Aerospace Exploration Agency
LE	Linear Error
LiDAR	Light Detection and Ranging
LightGBM	Light Gradient Boosting Machine
LLD	Land Levelling Datum
LMIC	Low- and Middle-Income Country
MAD	Median Absolute Deviation
MAE	Mean Absolute Error

MdAE	Median Absolute Error
ME	Mean Error
METI	Ministry of Economy, Trade and Industry of Japan
ML	Machine Learning
MLR	Multiple Linear Regression
NASA	National Aeronautics and Space Administration
nDSM	normalised Digital Surface Model
NGBoost	Natural Gradient Boosting
NMA	National Mapping Agency
NMAD	Normalised Median Absolute Deviation
PDP	Partial Dependence Plot
QGIS	Quantum GIS
RF	Random Forest
RGF	Regularized Greedy Forest
RMSE	Root Mean Square Error
RPC	Rational Polynomial Coefficients
RSIF	Remote Sensing Image Fusion
SAGA	System for Automated Geoscientific Analyses
SAR	Synthetic Aperture Radar
SD	Standard Deviation
SDG	Sustainable Development Goal
SHAP	SHapley Additive exPlanations
SPOT	Systeme Pour l'Observation de la Terre
SRTM	Shuttle Radar Topography Mission
TanDEM-X	TerraSAR-X add-on for Digital Elevation Measurements
TPI	Topographic Position Index
TRI	Terrain Ruggedness Index

TST	Terrain Surface Texture
UFP	Urban Footprint Processor
UTM	Universal Transverse Mercator
VHR	Very High Resolution
VRM	Vector Ruggedness Measure
WGS84	World Geodetic System 1984
XAI	eXplainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

LIST OF PUBLICATIONS

Published

Conference: Okolie, C.J., Mills, J.P., Adeleke, A.K. & Smit, J.L. (2023). Enhancing the vertical accuracy of Copernicus digital elevation model using tree-based machine learning models. *Abstr. Int. Cartogr. Assoc.*, 6, 193, https://doi.org/10.5194/ica-abs-6-193-2023.

Conference: Okolie, C.J., Mills, J.P., Adeleke, A.K., Smit, J.L., & Maduako, I. (2023). The explainability of gradient-boosted decision trees for digital elevation model (DEM) error prediction. Int. *Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-M-3-2023*, 161–168, https://doi.org/10.5194/isprs-archives-XLVIII-M-3-2023-161-2023 , 2023.

Dataset: Okolie, C.J. (2023). GBDT-based DEM enhancement: elevation, terrain and land cover maps. *Figshare repository*, Open licence CC BY 4.0. https://dx.doi.org/10.6084/m9.figshare.24346735.

Code: Okolie, C.J. (2023). Python codes for DEM error prediction using gradient-boosted decision trees, with default and Bayesian-optimised hyperparameters. Includes error metrics, explainability with PDP, SHAP and other extras. GitHub. MIT open-source licence. https://github.com/mrjohnokolie/dem-enhancement

Submitted/In review

Journal: Okolie, C.J., Mills, J.P., Adeleke, A.K., Smit, J.L., Peppa, M..V., Altunel, A.O., & Arungwa, I.D. Assessment of the global Copernicus, NASADEM, ASTER and AW3D digital elevation models in Central and Southern Africa. *Under review in Geo-spatial Information Science, Manuscript No. TGSI-2023-0082*.

Journal: Okolie, C.J., Adeleke, A.K., Mills, J.P., Smit, J.L., Maduako, I., Bagheri, H., Komar, T & Wang, S. Performance assessment of tree-based ensemble algorithms for the enhancement of Copernicus digital elevation model. *Submitted to the International Journal of Image and Data Fusion Geo-spatial information Science, Submission ID: 233007649.*

Conference: Okolie, C.J., Mills, J.P., Adeleke, A.K., & Smit, J.L. Digital elevation model correction in urban areas using extreme gradient boosting, land cover and terrain parameters. *Submitted to the 8th International Conference on GeoInformation Advances*. Arxiv preprint https://arxiv.org/abs/2308.06545.

Conference: Okolie, C.J., Ogbeta, C.O., Adeleke, A.K., Smit, J.L & Mills, J.P. Geo-morphometric comparison of global digital elevation models: NASADEM, Copernicus and ALOS World 3D. *Submitted to the 8th International Conference on GeoInformation Advances*.

To be submitted/In preparation

Journal: Okolie, C.J., Adeleke, A.K., Mills, J.P., Smit, J.L., Maduako, I. & Bagheri, H. An explainable gradient boosted decision tree-based feature-level fusion framework for the enhancement of Copernicus and AW3D digital elevation models. *In prep*

Journal: Okolie, C.J., Adeleke, A.K., Smit, J.L., & Mills, J.P. A statistical-based approach for correction of NASADEM in Cape Town, South Africa. *In prep.*

Journal: Okolie, C.J., Adeleke, A.K., Smit, J.L., & Mills, J.P. A systematic review of statistical and machine-learning based digital elevation model enhancement methods. *In prep.*

Conference: Okolie, C.J., Adeleke, A.K., Mills, J.P., Smit, J.L., Maduako, I. & Ogbeta, C.O. Computation time analysis of gradient boosted decision trees for digital elevation model (DEM) error prediction. *In prep.*

CHAPTER 1
INTRODUCTION

1.1 Background

For several decades, global digital elevation models (DEMs) have continued to gain traction in the geospatial, remote sensing and allied communities. Thanks to their analysis-ready formats and wide-area coverage, practitioners and mapping organisations have adopted global DEMs for numerous applications such as terrain analysis in geomorphology, production of topographic maps, hydrological modelling and analysis, mapping of natural hazards, urban planning and modelling, civil engineering and construction, modelling earth movements, rendering three dimensional (3D) visualisations, forest ecology, reduction of gravity measurements and the creation of digital twins. Unsurprisingly, the United Nations Committee of experts on Global Geospatial Information Management (UN-GGIM) have recognised DEMs within its "global fundamental geospatial data themes", under the category, "Elevation and Depth" (UN-GGIM, 2019). According to UN-GGIM (2019), elevation and depth are significant contributors to the attainment of the Sustainable Development Goals (SDGs). The utility and applications scope of DEMs are inextricably linked to their vertical accuracies, and it is important for DEMs to fulfil the quality requirements of their applications. Thus, the body of knowledge on accuracy assessment and enhancement of DEMs has experienced tremendous growth in recent years.

A DEM can be defined as a 2.5-dimensional (2.5D) representation of the surface of the earth in digital format, in which the elevation (z) is a bi-variate function, i.e., $z = f(x, y)$ (Polidori & Hage, 2020). The earth's surface is digitally and contiguously represented at a uniform pixel interval, and the pixel values correspond to elevation (World Bank Group, 2015). The basic information unit in a DEM is a point of height value *z*, and its corresponding *x* and *y* coordinates. The simple raster model (with regular matrices/uniform grids) is the most commonly used DEM format (Afana, 2011; Moore et al., 1991; Wilson & Gallant, 2000). Furthermore, the grid-based method is also efficient for analysing terrain derivatives and for spatial data manipulation (Aronoff, 1989; Jordan, 2007). DEMs are produced using several methods such as conventional ground surveys, remote sensing, and kinematic Global Navigation Satellite Systems (GNSS). Although airborne Light Detection And Ranging (LiDAR) is well recognised for its high quality, it is prohibitively expensive and used mostly in high-income countries (Meadows & Wilson, 2021). Space-borne LiDAR datasets are

gaining some traction but are still limited by extensive data gaps in their coverage. Satellite remote sensing exceeds the capabilities of airborne systems for acquiring and generating DEMs on a global scale. Today, there are several DEMs with global coverage available to users worldwide.

Global DEMs emanate from different acquisition techniques and differ in their accuracies, geometric properties, resolution characteristics and coverage (Okolie & Smit, 2022; Petrasova et al., 2017). Two of the most used methods for DEM generation from space-borne systems are Interferometric Synthetic Aperture Radar (InSAR) and photogrammetric techniques. The InSAR technique has progressed from theory to large scale applications by numerous research teams all over the world, and has been used in the production of several global DEMs Photogrammetric DEMs also enable remote mapping of global surface topography, especially with the emergence of sub-metre resolution satellite images and stereo-imaging capabilities (Noh & Howat, 2015).

The qualities of global DEMs are affected by several factors such as land cover variations and terrain anomalies. This compromises their reliability for environmental applications where accurate terrain information is required. Thus, researchers have adopted a variety of methods for enhancing their qualities. However, given the prohibitive financial expense of high-resolution topographic data and the emerging realisation of the importance of elevation data in global change and sustainable development research, the demand for freely available global DEMs has increased (Okolie & Smit, 2022). Their use and adoption have rapidly increased in recent times. These DEMs cover entire landmasses and are usually supplied digitally in raster format (Schindler et al., 2011). Geospatial and geomatics practitioners in Africa have been combining conventional topographic data acquisition sources with global DEMs for revision and updating of topographic maps (e.g. Owino, 2021; Pepple et al., 2023; Nwilo et al., 2017; Nwilo et al. 2022a; Olusina & Okolie 2018).

Several methods have been proposed for correcting DEM errors, e.g., multiple linear regression (Su et al., 2015), Bayesian information criterion (Zhou et al., 2020), subtraction of vegetation heights (Baugh et al., 2013), multivariate linear regression using vegetation and slope information (Su & Guo, 2014), the use of power models (O'Loughlin et al., 2016), spherical harmonics (Wendleder et al., 2016), interpolation of error surfaces (Zhao et al., 2015), and machine learning algorithms (Hawker et al., 2022). Generally, the approaches are grouped into manual techniques and systematic approaches. The manual approach involves direct editing of

a DEM using expert judgement or supplementary data, while systematic editing uses explicit algorithms that could include ancillary datasets or filters to screen out errors in the DEMs (Hawker et al., 2018; Meadows & Wilson, 2021). However, methodological frameworks that can integrate a suite of features such as terrain parameters (e.g., slope and aspect) and land cover into an effective, robust, coherent and adaptive system for DEM enhancement are required. Also, given the fact that many mapping agencies or organisations, scientists and researchers in Africa are heavily reliant on global DEMs, a clear guidance for their enhancement is required. This framework when implemented could yield reliable and more comprehensive elevation information for improved flood/erosion prediction, mapping of tropical forest biomes, determining the 3D vertical structure of savannah and dense forests; and for hydrological and water resources management.

Machine learning is well suited for DEM enhancement because it can uncover non-linear and complex relationships between variables, and handle large volumes of remote sensing data across extensive swathes of land (Kulp & Strauss, 2018; Lary et al., 2018; Meadows & Wilson, 2021). This research is also inspired by tree-based ensemble machine learning algorithms which are very reliable and widely adopted (Papadopoulos et al., 2017). These models present several advantages such as interpretability, less data preparation, versatility, and ability to handle non-linear and complex relationships. Decision trees are easy to understand (Miao et al., 2011), their logical rules are easier to interpret (Kotsiantis, 2013), and they are highly tolerant to multicollinearity (Climent et al., 2019; Han et al., 2019; Pham & Ho, 2021). Moreover, Gradient Boosted Decision Trees (GBDTs) have emerged as a winning solution in several data mining and machine learning tasks (e.g., Hutchinson et al., 2011; Natekin & Knoll, 2013; Pittman & Brown, 2011) and in several studies on tabular datasets, have outperformed deep learning algorithms (e.g., Kadra et al., 2021; Borisov et al., 2022; Shwartz-Ziv & Armon, 2022). Despite these modest achievements, the full potential of tree-based algorithms is yet to be fully tested or materialised in the field of remote sensing, and specifically, for DEM enhancement. To the best of this author's knowledge, GBDTs have never been applied for DEM enhancement in any other published research.

1.2 Statement of the Problem

Generally, the feature-based strategy for DEM enhancement involves the fusion of datasets (including terrain and land cover parameters) for prediction of the elevations, or the elevation error which is subsequently applied for correcting the DEMs. However, to keep pace with the

growing need for digital elevation data and topographic information, new processing paradigms are needed to automate the enhancement of global DEMs. Some important terrain parameters were not considered in previous approaches, and the diversity and complexity of variable landscapes have not been fully addressed. Some of the reviewed studies which were based on machine learning (ML) are individual or separate ML implementations, without high-level optimisation. From the available evidence, the implementation of a feature-based DEM enhancement strategy which could exploit the complimentary predictions of multiple tree-based machine learning ensembles has never been attempted. Therefore, a methodology that exploits the relationships and interactions between terrain and land cover parameters, and the complementarity of tree-based ensemble learners into a unified framework for DEM enhancement will amount to a significant advancement in the state-of-the-art.

Going further, the present study incorporates terrain characteristics from a wide variety of landscapes that cut across urban/industrial, agricultural, mountain, peninsula and grassland/shrubland landscapes. This is for a more robust solution with relevance in different landscape contexts. The enhancement strategy developed will establish a methodology that is applicable to a wide variety of digital elevation datasets and suitable for deployment in various climes. Finally, the use of restricted datasets and unexplainable (or uninterpretable) machine learning techniques in remote sensing studies have confounded their adoption and reproducibility by practitioners. Model explainability is becoming a major requirement for deployment of machine learning models. Despite their high-level performance, many state-of-the-art models are viewed as black boxes which are difficult to explain or interpret (Kim, 2021). This lack of explainability of many complex ML models is a source of tension between accuracy and interpretability (Lundberg & Lee, 2017). With regard to ML-based DEM enhancement approaches, there are several unknowns regarding the interdependencies between terrain parameters and their influences on the predicted elevation error. Thus, this research prioritises an explainable ML-based DEM enhancement workflow for rapid adoption by end-users.

1.3 Aim and Objectives

The aim of this research is the enhancement of satellite-derived multi-sensor (interferometric and photogrammetric) global DEMs through the implementation of a tree-based ensemble feature-level fusion scheme that exploits the interactions between terrain and land cover parameters. The specific objectives are:

1. To carry out a comprehensive quality assessment of specific interferometric and photogrammetric global DEMs to inform their selection for further analysis.

2. To align the selected DEMs, and integrate a collection of extracted terrain and land cover parameters for characterisation of the target landscapes.

3. To develop a data fusion methodology for DEM enhancement by modelling the relationship between the terrain/land cover parameters and elevation error using an assemblage of default and Bayesian hyperparameter-optimised tree-based ensembles.

4. To interrogate the interdependencies/interactions between the terrain/land cover parameters, and their influence on the predicted elevation errors.

5. To evaluate the performance of the DEM enhancement scheme and compare the results with statistical-based DEM correction, alternative DEMs and results from previous studies.

1.4 Research Questions

The following research questions are investigated:

1. How does the quality of the selected DEMs differ in various landscapes?
2. What are the relationships between terrain parameters and the vertical accuracy?
3. Can the accuracy of elevation error predictions be significantly improved through tree-based ensemble ML feature-level fusion?
4. What are the magnitudes and directional impacts of the interdependencies/interactions between terrain and land cover parameters in tree-based ensemble learning for DEM error prediction?

1.5 Scope of Research

Qualitative and quantitative analytical techniques are adopted for assessing the qualities of the four global DEMs using high-resolution airborne LiDAR as a reference at different landscapes in Cape Town. Following this assessment, terrain and land cover parameters are integrated into a tree-based ensemble ML feature-level fusion scheme for predicting the elevation error distribution. The learning process is further optimised through efficient Bayesian optimisation with an extensive search space designed to yield hyperparameters in an informed manner and return optimal predictions. The predicted elevation errors are applied for correcting the DEMs

at several model implementation sites. Finally, performance evaluations and comparative assessments are conducted to validate the proposed enhancement scheme.

1.6 Significance of Research

Enhanced DEMs with improved vertical accuracies provide a much more hydrologically correct elevation database for environmental modelling in applications such as groundwater potentiality mapping, watershed/catchment delineation, modelling of surface geo-morphometric parameters, flood vulnerability modelling, and erosion risk assessment. There is also the potential for DEM enhancement methodologies to be adapted in the landscape contexts of other countries. Enhanced DEMs can be used for models that show how water is stored, transported and used, and the derived terrain parameters can also provide the physical environmental information needed to map ecosystems and species distributions. From enhanced DEMs, a suite of commonly used terrain attributes can be derived, which have wide applications for ecological, hydrological and geomorphological processes. More importantly, this study presents a novel DEM enhancement methodology as an impetus for mapping agencies to implement deliberate measures aimed at ensuring more comprehensive and accurate updates of their national topographic databases as a critical component of geospatial data infrastructure.

1.7 Contributions to Knowledge

Several novel contributions to knowledge are made in this research which presents an innovative approach for deriving highly accurate and precise elevation data in urban/industrial, agricultural, mountain, peninsula and grassland/shrubland landscapes, from open-access global DEMs.

To the best of this author's knowledge:

1. This is the first research to perform a comparative assessment of bagging versus boosting tree-based ensemble machine learning approaches applied to DEM enhancement in Africa. Moreover, the sheer number of tree-based ensemble models included in the comparative analysis provides a significant addition to the body of knowledge on their relative performances for handling remote sensing tasks and problems in other research fields of endeavour.
2. This is the first comparative assessment of the selected number of gradient boosted decision trees (GBDTs) for DEM enhancement.

3. This is the first time that natural gradient boosting (a newly developed and innovative probabilistic regression algorithm which utilises natural gradients instead of regular gradients) is tested for returning point estimates of elevation error.
4. This is likely the first research to integrate model explainability measures for understanding the magnitudes and directional impacts of variable interactions in tree-based ensemble learning for DEM error prediction and correction. Moreover, this research improves the overall understanding of the spatial characteristics of DEM error.
5. This is the first machine learning-based enhancement of satellite-derived interferometric and photogrammetric open-access global DEMs implemented over field sites in Africa.

Moreover:

6. The corrections are executed on a pointwise basis, thus reducing systematic errors at each grid/cell in the DEM.
7. The inexpensive data requirements of the proposed DEM enhancement scheme enable its application to other landscapes for which hydrological modelling accuracy is hindered by terrain offsets.

Finally, the proposed DEM enhancement scheme:

8. Is scalable, adaptable, efficient and a low-cost methodology for improving the quality (and accuracy) of DEMs.
9. Is capable of learning complex and non-linear terrain information, and multi-variate spatial patterns in areas with groundtruth, with subsequent application for DEM error prediction and correction at other areas where groundtruth data is non-existent.
10. Incurs ‘minimal’ computational expense while delivering competitive improvements in accuracy.
11. Has the potential for the improvement of terrain representation, diminution of terrain and above-ground offsets, and for the generation of bare-earth DEMs or digital terrain models (DTMs) in urban and tree-covered areas.

1.8 Thesis Outline

The introductory sections in Chapter 1 have provided a background and context for this research. Chapter 2 presents a comprehensive review of relevant literature regarding DEMs

and terrain parameters, quality assessment of DEMs, machine learning techniques for DEM enhancement and other related topics. Chapter 3 provides a foundational review of the theories that serve as core knowledge for the research. Chapter 4 elaborates on the methodological framework, study area and datasets, the processing workflows, the implementation stages of the DEM enhancement framework, performance evaluation and comparative assessments.

Chapter 5 provides an extensive analysis and discussion of the results from the DEM quality assessment, machine learning and statistical-based corrections, and the performance evaluations and comparative assessments. In Chapter 6, conclusions and recommendations are made, based on the performed analysis.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

This chapter provides a detailed understanding of relevant academic literature about topics within the context of this research, specifically: digital elevation models, terrain parameters, remote sensing data fusion with applications to DEM enhancement, and tree-based ensemble machine learning techniques.

2.2 Digital Elevation Models

Generally, a Digital Elevation Model (DEM) refers to digital topographic data which can be analysed and modelled by a computer in three dimensions in a manner that reduces the requirement for labour-intensive human interpretation (Maune, 2001). Guth et al. (2021, p. 7) define a DEM as "a digital representation of elevations (or height) of a topographic surface in form of a geo-rectified point-based or area-based grid, covering the Earth or other solid celestial bodies". An earlier definition by Moore et al. (1991) states that DEMs are: "an ordered array of numbers that represent the spatial distribution of elevations above some arbitrary datum in a landscape". The basic information unit in a DEM is a point of height value z, and its corresponding *x* and *y* coordinates. According to Washburn (2017), the modern DEM concept gained popularity in the 1960s.

The advent of the digital computer in the 1950s brought advancement in the automation of techniques for topographic surface modelling. Continuous progress in digital computing led to more sophisticated DEM generation and analysis methods, and the proliferation of desktop computers in the 1990s led to more complex analysis in spatial software packages (e.g., Pike et al., 2009; Washburn 2017). Today, DEMs are a core component of Geographic Information Systems (GIS) and have found wide applications in environmental analysis, geomorphology, engineering, ecology and hydrology (Jordan, 2007). Also, with new developments in terrain reconstruction technologies, topographic data is becoming increasingly available in a wide range of formats (Petrasova et al., 2017).

Applications of DEMs include hydrological simulations (Rocha et al., 2020; Xu et al., 2021), flood hazard mapping (Leenaers & Okx, 1989), cartographic relief presentation (Samsonov, 2022), infrastructure surveys and mapping (Barbarella et al., 2022; Van Nieuwenhuizen et al.,

2021), archaeology (Štular et al., 2021), land cover mapping (Ibrahim, 2023) and soil erosion analysis (Michalopoulou et al., 2022).

In the existing literature, some other terminologies are closely related to 'DEM' such as (Guth et al., 2021; Podobnikar, 2005):

1. **Digital surface model (DSM)**: A DSM captures the lower boundary of the atmosphere (and either the lithosphere, anthroposphere, cryosphere, hydrosphere, or biosphere). Many of the freely available global elevation datasets fall under the category of DSMs.
2. **Digital terrain model**: A DTM captures the boundary between the lithosphere (the solid outer part of the earth) and the atmosphere, without anthroposphere and biosphere. A DTM can include the values of elevation and other terrain elements (e.g., slope).
3. **Normalised DSM (nDSM):** The difference between a DSM and DTM is referred to as a Normalised DSM. It represents the elevations of features on the earth's surface including trees and buildings.

The terrain representation by DTM and DSM is shown in Figure 2.1. Guth et al. (2021) have compiled a comprehensive glossary of terminologies and definitions related to DEMs. DEMs are usually prepared using two main data models/structures: raster models with hierarchical or regular matrices (Afana, 2011; Burrough & McDonnell, 1998), and vector models such as contour lines and triangulated irregular networks (Afana, 2011; Peucker et al., 1978). Due to their ease of implementation and simplicity, DEMs with uniform matrices or regular grids are the most common structure type (Afana, 2011; Moore et al., 1991; Wilson & Gallant, 2000).

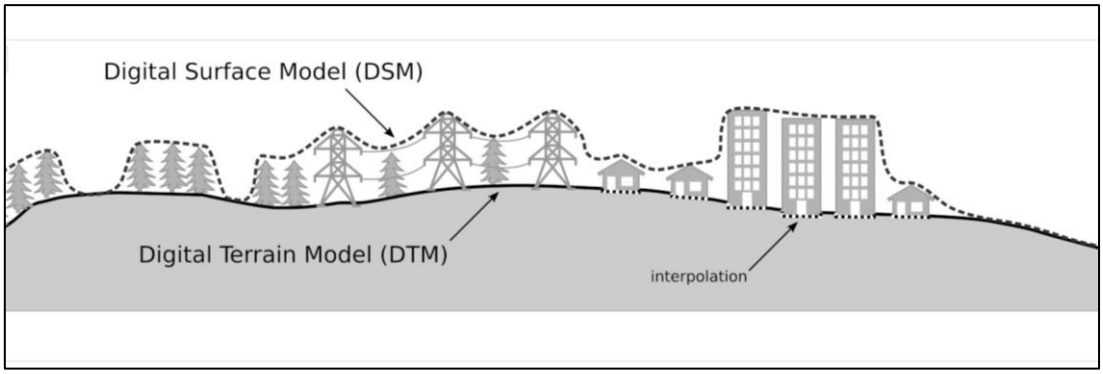


Figure 2.1: Terrain representation by DSM and DTM (Adapted from Guth et al., 2021, p. 8)

Airborne and spaceborne remote sensing systems are complementary to traditional field surveying and photogrammetric approaches for generating DEMs. However, when sensing the earth from air or space, data is collected from the first surface with which the electromagnetic

energy interacts with. The resulting surface models are representative of the elevation of that terrain composed of both the "bare Earth" surface and above-ground objects or features (shown in Figure 2.2). Although this first surface is extremely useful for diverse applications involving landscape analysis, it does not provide the true-earth elevation data required by several users for detecting the bare Earth surface (Hu, 2003).

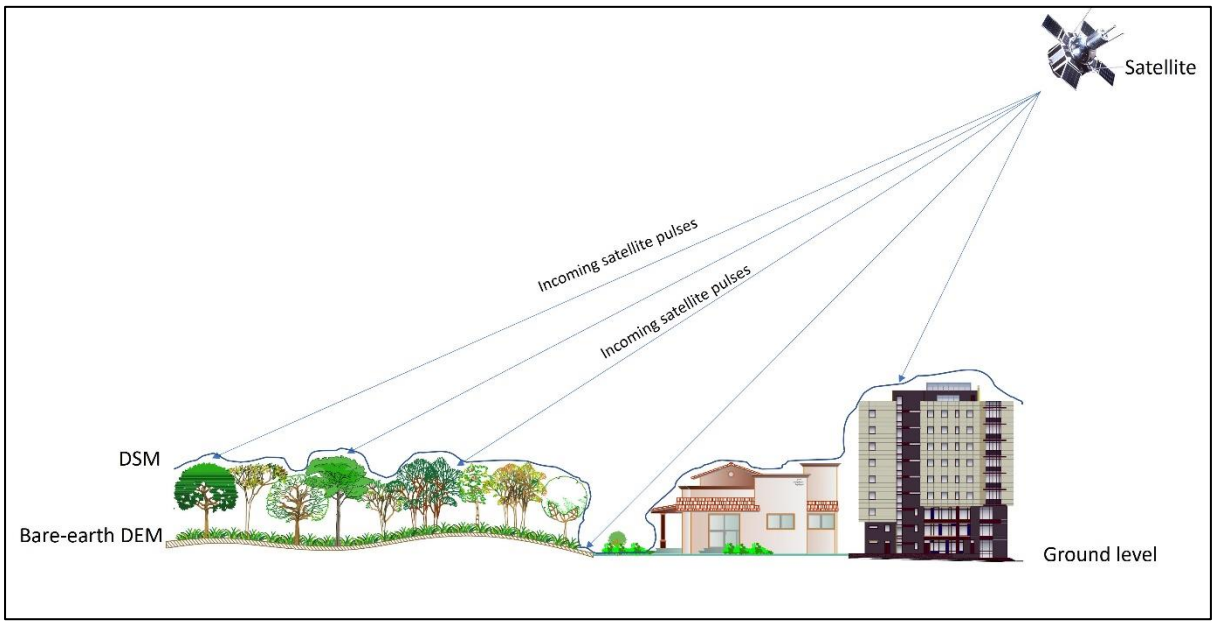


Figure 2.2: The influence of above-ground features which obscure the topography of the bare-earth (Source: Okolie & Smit, 2022, p. 8)

The bare-earth is a continuous surface with nothing visible below it and is usually represented by the DTM in the civil engineering and surveying engineering fields (Haugerud & Harding, 2001; Hu, 2003). While the bare-earth consists of the topsoil and thin pavements, the above-ground features are composed of detached objects (e.g., building and trees) and attached objects (e.g., bridges and viaducts) respectively (Hu, 2003). A bare-earth DEM provides an opportunity to quantitatively characterise the land surface and yield digital terrain information without the blur caused by land cover features (see Figure 2.2).

2.2.1 Global DEMs

Global DEMs represent "homogeneous topography information for nearly all landmasses of the world" (Bagheri et al., 2018a). Global DEMs are produced using satellite remote sensing technology, and provide global or near-global elevation coverage. Airborne sensors are difficult to scale up for global coverage because the data capture costs are prohibitive. Due to the open data policy of many national space and earth observation agencies, global DEMs are increasingly accessible to scientific end-users.

2.2.2 National DEMs

National DEM can be regarded as primary elevation datasets covering a country, produced and/or distributed by its national mapping agency (NMA) or a delegated government authority. National DEMs are designed to provide elevation data of the territorial extent of a country in a seamless format and with consistencies in elevation units, datum and projection (Papasaika-Hanusch, 2012), and are generally expected to be more detailed and accurate than global DEMs. Methods for producing national DEMs include traditional photogrammetry and LiDAR. According to Yuan et al. (2009), LiDAR is becoming the prime method for large scale acquisition of nationwide DEMs. However, Very High Resolution (VHR) DEMs produced by airborne LiDAR systems are expensive and hence, not widely available (Baugh et al., 2013; Pham et al., 2018). Examples of DEMs with national coverage include OS Terrain 5 (Great Britain; Ordnance Survey, 2023), DTED-2 (Turkey; Yilmaz & Erdoğan, 2018), Tinitaly DEM (Italy; Tarquini et al., 2012), RGE ALTI (France; IGN, 2018), HRDEM (Canada; Government of Canada, 2023) and the South Africa National DEM (South Africa; CD: NGI, 2019).

2.3 Methods for Producing Global DEMs

Some common methods for producing global DEMs are discussed below. These techniques are well documented in existing works and so, only a brief overview is presented here.

2.3.1 Synthetic aperture radar interferometry

Synthetic Aperture Radar (SAR) from satellite platforms has shown enormous potential for DEM generation (e.g., Agrawal et al., 2018). DEMs can be generated to metre-scale accuracy from synthetic aperture radar (SAR) imagery through the process of interferometry. This method is referred to as synthetic aperture radar interferometry (InSAR). InSAR is the analysis of two SAR images captured over the same area under slightly varying incidence angles (Hoja et al., 2006). It has progressed from theory to large scale applications by numerous research teams all over the world. It is now a scientifically accepted methodology with wide application areas (Smith, 2002; Geymen, 2014). It has been applied in the generation of global DEMs such as the Shuttle Radar Topography Mission (SRTM). InSAR has also been applied to the TerraSAR-X add-on for Digital Elevation Measurements (TanDEM-X) shown in Figure 2.3 (Krieger et al., 2007, 2013).

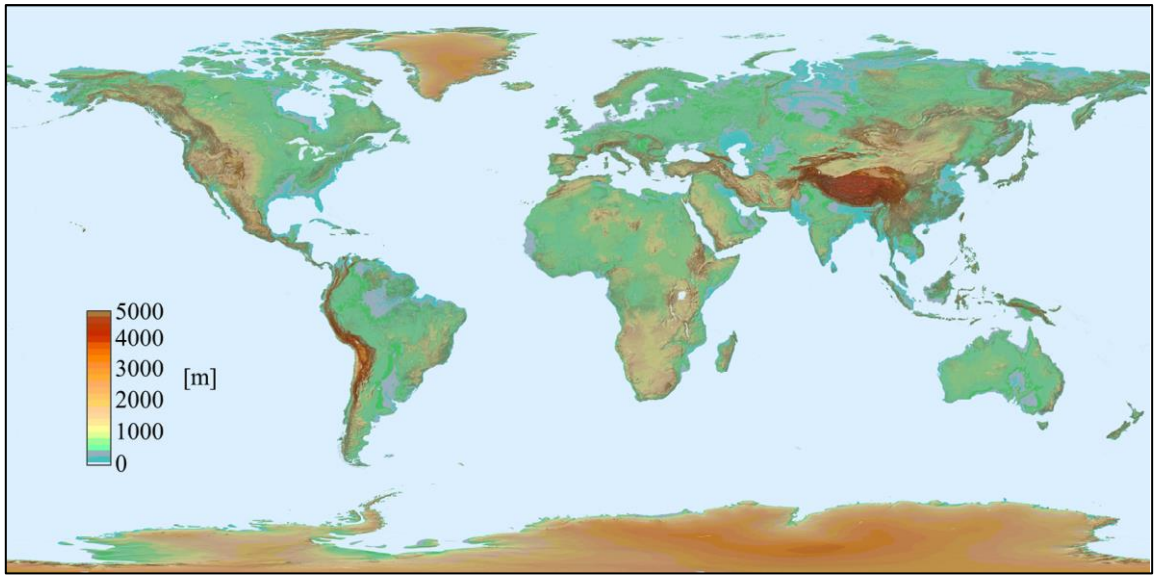


Figure 2.3: Visualisation of the global TanDEM-X DEM (Source: Rizzoli et al., 2017)

The advantages of using SAR/InSAR for DEM generation include the all-weather and full day-and-night capability, sensitivity of SAR signal to the earth's surface, and robust and analytical techniques (Feng et al., 2020; Mohammadi et al., 2020; Xiang et al., 2017; Pepe & Calò, 2017). InSAR techniques for DEM generation are well documented (e.g., Gens & Genderen, 1996; Braun, 2021; Yang et al., 2020).

2.3.2 Radargrammetry

The first studies to use radargrammetry were conducted in the 1950s. Unlike interferometry, radargrammetry is less affected by atmospheric influences because it uses the amplitude images which are less affected by atmospheric disturbances (Yu et al., 2014). It is commonly applied in SAR images with varying incidence angles and can also be applied in images with varying aspect angles (Feng et al., 2020). The image pairs with different incidence angles can be in the same node or in opposite nodes with a short time interval (Agrawal et al., 2018). Radargrammetry is increasingly being applied to DEM generation thanks to the emergence of VHR SAR satellites such as COSMO-SkyMed, RADARSAT-2 and TerraSAR-X (Capaldo et al., 2011; Papasodoro et al., 2016; Raggam et al., 2010; Toutin, 2010, 2012). A key advantage of radargrammetry is its capability to rapidly generate DEMs without the need for ground control points (GCPs), thanks to the availability of accurate geometric models, orbit information and rational polynomial coefficients (RPCs) (Papasodoro et al., 2016).

2.3.3 Stereo-photogrammetry

The basic principle of producing DEMs from satellite stereo imagery is to match the feature points based on the image overlaps, the generation of the three-dimensional (3D) coordinates

from the feature points, and the reconstruction of a spatial stereo model and obtaining of the ground 3D information (Wang et al., 2019). Since the launch of the Systeme Pour l'Observation de la Terre (SPOT) satellite images since 1986, DEM generation via stereo images has seen many innovations and advancements (Yanalak et al., 2012). Different satellite sensors have been launched with stereo imaging capability and several photogrammetrically derived DEMs have been released, e.g., the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global DEM shown in Figure 2.4, and the Advanced Land Observing Satellite (ALOS World 3D) DEM.

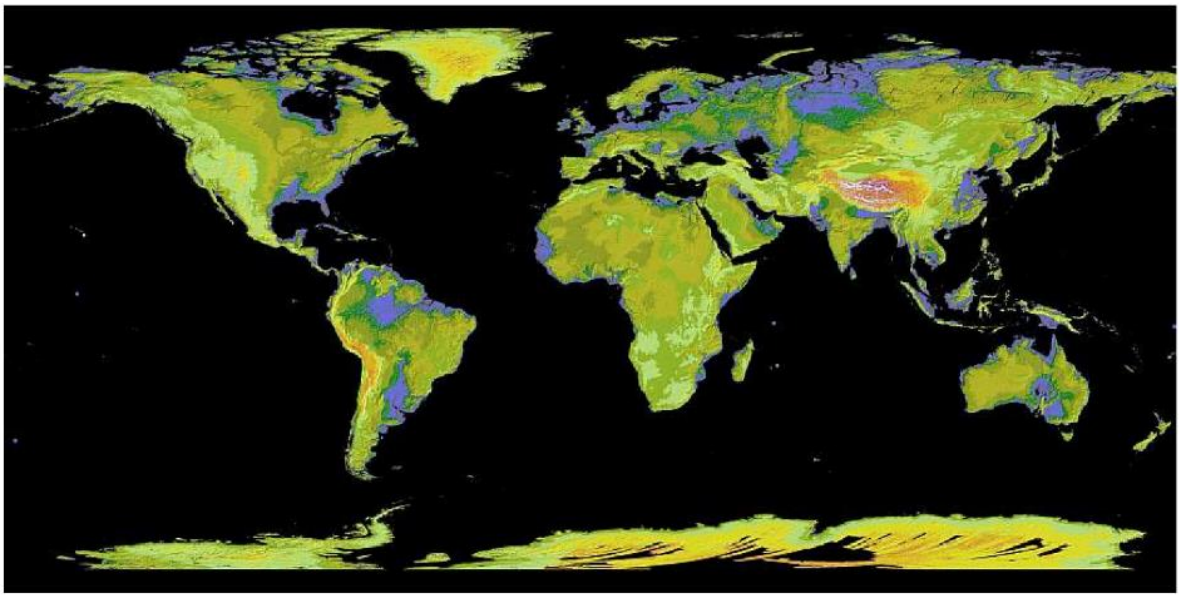


Figure 2.4: View of ASTER version 1 coverage (Source: Abrams et al., 2020)

2.3.4 Light detection and ranging

The integration of Light Detection And Ranging (LiDAR) sensors onboard earth observation satellites has ushered in a new generation of global DEMs (Vernimmen & Hooijer, 2023). These LiDAR sensors are known to provide higher vertical accuracies than earlier radar data sources. However, they do not provide very detailed ground coverage due to limitations in the LiDAR footprint density. The Geoscience Laser Altimeter System (GLAS) aboard the Ice, Cloud and Land Elevation Satellite (ICESat) is a LIDAR sensor that provides high resolution measurements of ground and vegetation heights (Pang et al., 2008). More recent satellite LiDAR missions include ICESat-2 (Markus et al., 2017) and the Global Ecosystem Dynamics Investigation, GEDI (Dubayah et al., 2020). Figure 2.5 and Figure 2.6 show ICESat and GEDI global elevation maps respectively.

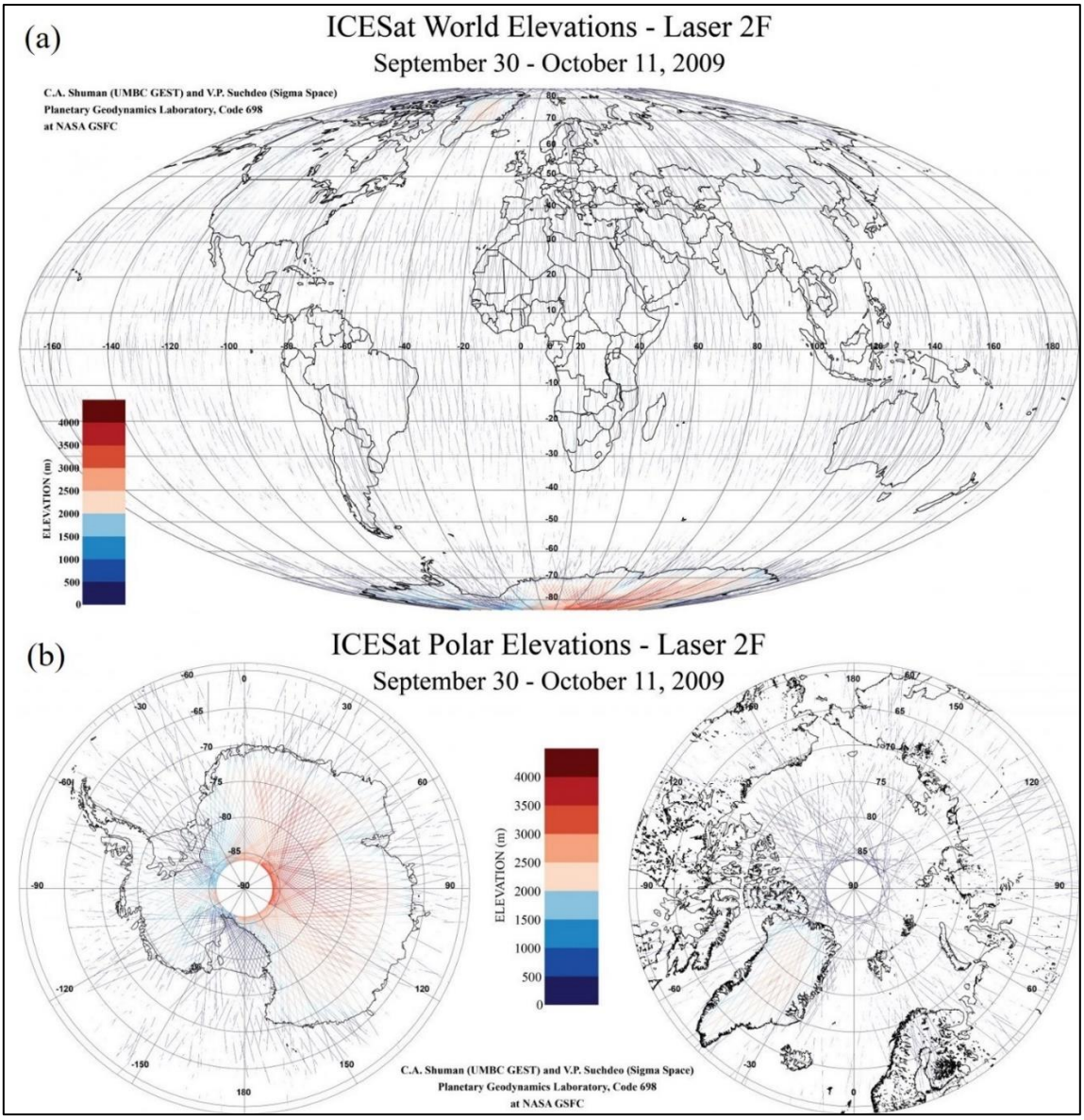


Figure 2.5: ICESat (a) World and (b) Polar elevation maps (Image courtesy of C.A Shuman and V.P. Suchdeo, NASA Goddard Space Flight Centre)

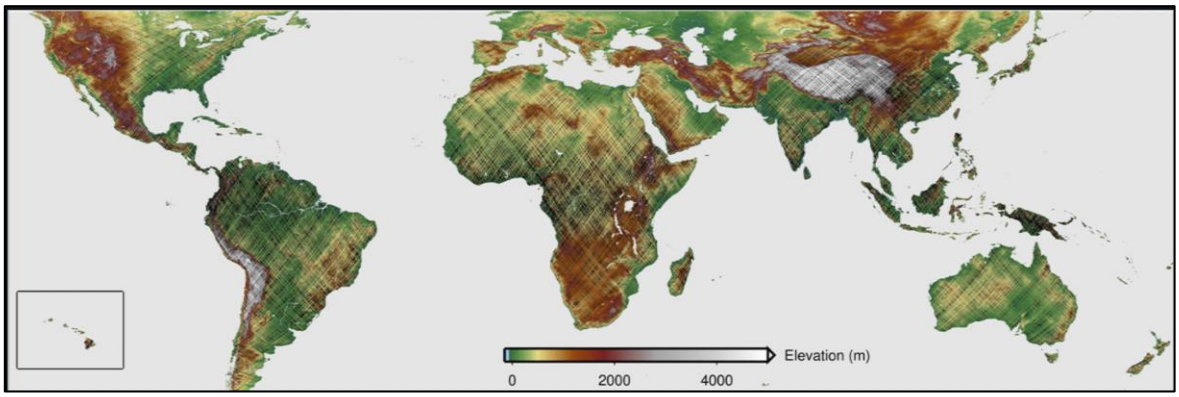


Figure 2.6: Global 'bare-earth' elevation map from the GEDI mission (Image courtesy of NASA Earthdata)

The accuracies of global DEMs are influenced by a multitude of factors. Thus, assessing the qualities of DEMs is pertinent for a holistic perspective on their utility and performance limitations.

2.4 Quality Assessment of DEMs

Quality assessments are required to detect and understand the effect of errors on the vertical accuracy of DEMs. Also, quality assessment is required to determine if a DEM fulfils certain user requirements that are based on adapted quality criteria (Polidori & Hage, 2020; Reuter et al., 2009). Errors in DEMs can be categorised into three types shown in Figure 2.7 (Papasaika-Hanusch, 2012; Polidori & Hage, 2020; USGS, 1997):

1. **Systematic errors:** systematic errors extend across large areas or regions and can be detected through expert analysis of the terrain, or by comparison of the DEM with existing data of higher accuracy. Systematic errors exhibit a fixed trend or pattern, are usually of constant magnitude and are predictable.
2. **Blunders:** Blunders are vertical errors such as spikes that are usually of a high magnitude and that exceed the maximum absolute error. Methods of blunder detection include DEM visualisation and methods based on anomalous slopes.
3. **Random errors:** Random errors are not easily detected by DEM users. They include mistakes caused by inaccurate surveying or erroneous recording of elevation information.

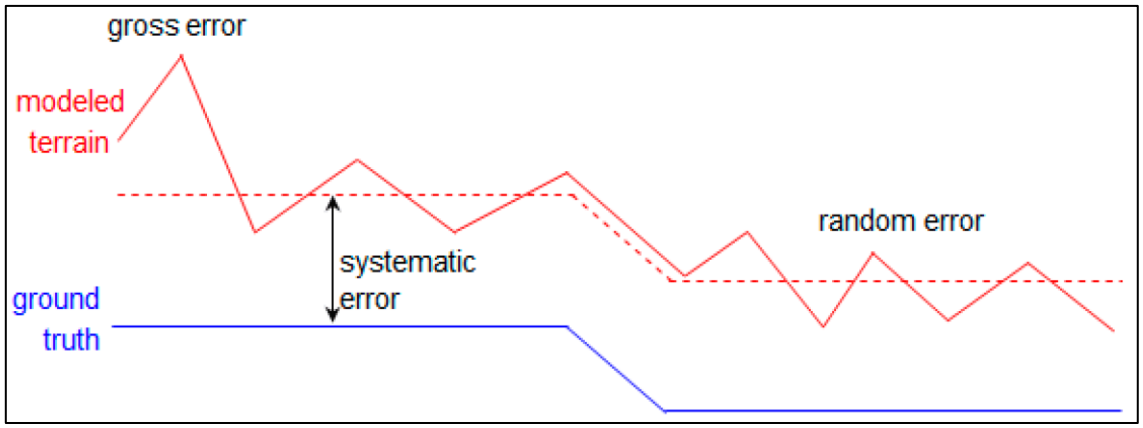


Figure 2.7: Three types of DEM error (Source: Polidori & Hage, 2020)

Qualitative (visual) and quantitative analysis are among the standard and advanced checks described in Gesch et al. (2014) and Cenci et al. (2021) as quality assessment criteria for DEMs. There are two main DEM quality assessment approaches (Polidori & Hage, 2020):

1. **External validation:** in external validation, the quality of the DEM is assessed using external reference data such as ground control points (GCPs), contour lines, topographic profiles, LiDAR point clouds or a much more accurate DEM.
2. **Internal validation:** this type of quality assessment is implemented without ground control. Inconsistencies are detected in the DEM based on the criteria of realism and a priori knowledge of the general nature of topographic surfaces. Visual analysis is a basic and powerful approach that can quickly reveal the existence of problems in DEMs (Mesa-Mingorance & Ariza-López, 2020). To reduce subjectivity in the visual interpretation, experts and trained analysts are required to perform it (Cenci et al., 2021).

Vertical accuracy is an important quality specification for elevation data (ASPRS, 2004; Mesa-Mingorance & Ariza-López, 2020). Vertical accuracy is defined as "the measure of the positional accuracy of a data set with respect to a specified vertical datum, at a specified confidence level or percentile" (ASPRS, 2014, p. a5). Usually, the absolute vertical accuracy is assessed by calculating the pixel-by-pixel difference between the satellite DEMs and the reference DEM (LiDAR) (Cenci et al., 2021). The analysis of the elevation difference enables the detection of the presence of possible random errors and systematic bias (Cenci et al., 2021). Several reviews of DEM quality assessment methods have been published (e.g., Mesa-Mingorance & Ariza-López, 2020; Polidori & Hage, 2020).

2.5 Terrain Parameters derived from DEMs

Terrain parameters (or topographic attributes) reveal statistically significant differences on a topographic surface, and enable the modelling of the spatial variability of certain landscape processes (Gallant & Wilson, 2000).

2.5.1 Slope and aspect

Slope and aspect are the first derivatives of the mathematically continuous surface of a gridded DEM (Kienzle, 2004). They describe the structure and relief of the topographic surface and are important variables in landslide monitoring, mass movement studies and land use planning (Zhou & Liu, 2008). They are also important parameters for describing the terrain structure and morphology, surface process simulation, and hydrological and land use modelling (Florinsky, 2010; Chow & Hodgson, 2009; Gao, 2010; Wise, 2007; Zhu et al., 2019; Ziadat, 2007).

The accuracy of DEM-derived slope and aspect depends on the terrain complexity, the DEM error and resolution, and the processing algorithm (Gao, 2010; Chang & Tsai, 2013; Zhu et al., 2019). Given a point on the terrain defined by a function $z = f(x, y)$, slope (S) and aspect (A) can be expressed as follows (Zhou & Liu, 2008).

$$S = \arctan\sqrt{p^2 + q^2} \tag{2.1}$$

$$A = 180^{\circ} - \arctan\left(\frac{q}{p}\right) + 90^{\circ}\left(\frac{p}{|p|}\right) \tag{2.2}$$

Where,

p is the gradient at W-E direction

$$p = f_x = \frac{\partial f}{\partial x} \tag{2.3}$$

q is the gradient at N-S direction

$$q = f_y = \frac{\partial f}{\partial y} \tag{2.4}$$

Several methods have been developed to derive p and q, and are summarised in Zhou & Liu (2008).

2.5.2 Surface roughness

The concept of 'roughness' permeates the Earth Sciences, and is important for studying earth surface processes, identifying landforms and for describing topography (Lane 2005; Smith 2014; Grohmann et al., 2011). Generally, the terrain roughness measures the level of variation of the z-axis across the sampled area of terrain, within a defined area at a defined scale (Department of Environment and Science Queensland, 2020). It is used in earth sciences for delineating landform components and for inferring process information to derive geomorphological units (Department of Environment and Science Queensland, 2020). Roughness could also refer to the irregularity or variability of elevation, or even the elevation range (e.g. Moser et al., 2007; Wolf et al., 2011; Smith, 2014). Factors that influence the surface roughness of a landform include its age, material properties, and the nature of processes acting upon it (Grohmann et al., 2011).

The expected relationship between surface roughness measurements and the DEM spatial resolution is shown in Figure 2.8.

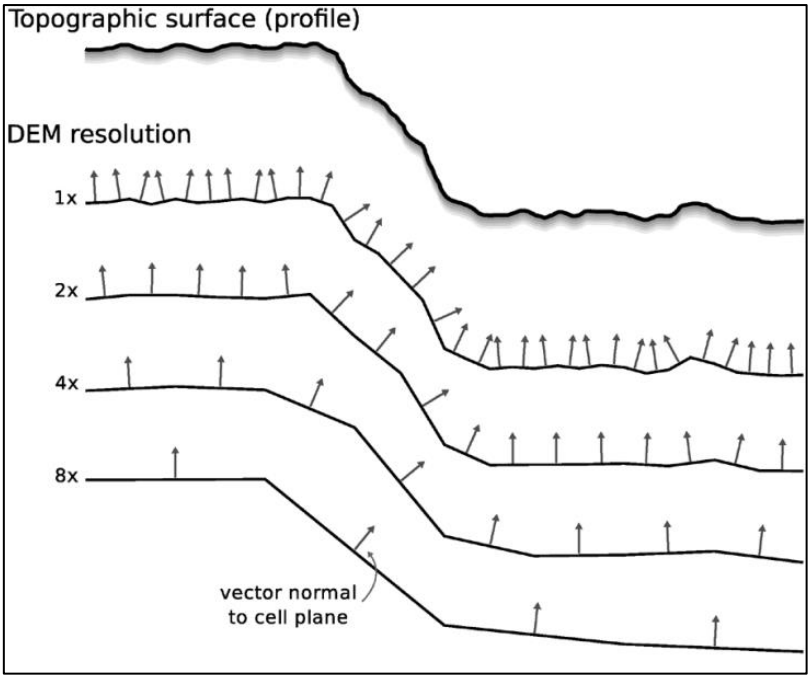


Figure 2.8: The variation of surface roughness according to DEM spatial resolution (Source: Grohmann et al., 2011, p. 1202)

2.5.3 Topographic position index

The Topographic Position Index (TPI) has applications in the measurement of slope positions, and for morphological and landform classification (De Reu et al., 2013). The expression for the TPI is derived by subtracting the average elevation ($\bar{z}$) around a central point (z_0) from the elevation of that central point, within a predetermined radius (R) (De Reu et al., 2013; Gallant & Wilson, 2000; Weiss, 2001):

$$TPI = z_0 - \bar{z} \tag{2.5}$$

$$\bar{z} = \frac{1}{n_R} \sum_{i \in R} z_i \tag{2.6}$$

A positive TPI means that z_0 is located higher than its average surroundings, and vice versa (De Reu et al., 2013).

2.5.4 Terrain ruggedness index

The Terrain Ruggedness Index (TRI) expresses "the amount of elevation difference between adjacent cells of a DEM" (USNA, 2022, para. 1). The elevation difference from a centre cell and the immediate eight surrounding cells are calculated; and then the eight elevation difference values are squared, summed, and the square roots are calculated (USNA, 2022). When interpreting TRI values, potential biases in the DEMs should be taken into consideration (Riley et al., 1999).

2.5.5 Terrain surface texture

Terrain Surface Texture (TST) was proposed by Iwahashi & Pike (2007) for terrain classification. Summarily, it is calculated by smoothening the input elevation with a median filter, extracting pits and peaks, and calculating the spatial frequency of pits and peaks (Iwahashi & Pike, 2007; GRASS GIS Manual, 2023).

2.5.6 Vector ruggedness measure

The Vector Ruggedness Measure (VRM) is a measure of terrain ruggedness. It is the variation in the 3D orientation of elevation pixels in a defined neighbourhood (Welty & Jeffries, 2018). The VRM was introduced by Sappington et al. (2007) and is based on the earlier method proposed by Hobson (2019) for measuring surface roughness. Because the VRM measures terrain variation independent of its overall gradient, it can differentiate among terrain types (Sappington et al., 2007). Moreover, VRM enables the treatment of terrain components as separate variables when quantifying landscapes (Sappington et al., 2007; Welty & Jeffries, 2018). Ruggedness values can range from 0 to 1. The equation for the VRM is given by Pipaud et al. (2015).

2.6 Remote Sensing Data Fusion

Data fusion is "a common paradigm related to the processing of data observed by different sensors and finds its place in a large variety of domains" (Mura et al., 2015, p. 1586). The methods of data fusion originated from the military, but have since found wide applications in civilian circles (Hall & Llinas, 1997). According to Sidek & Quadri (2012, p. 3), "data fusion deals with the synergistic combination of information made available by different measurement sensors, information sources, and decision makers." Similarly, Schmitt & Zhu (2016, p. 8) opine that "data fusion is all about combining the observations from more than one sensor, position, or time (and ideally even from more than one sensor type) to infer unknown parameters about the object of interest." Fusion is made possible because each dataset presents unique and different views of the same real world, with each view having its own sampling procedures, measurement processes and measurable properties (Ghamisi et al., 2019). It is therefore expected that there should be some level of coherence between the datasets (or observations) and the real world (the source), and this concept of coherence is fundamental to data fusion (Ghamisi et al., 2019; Wang et al., 2015). Figure 2.9 shows the generalised data fusion process.

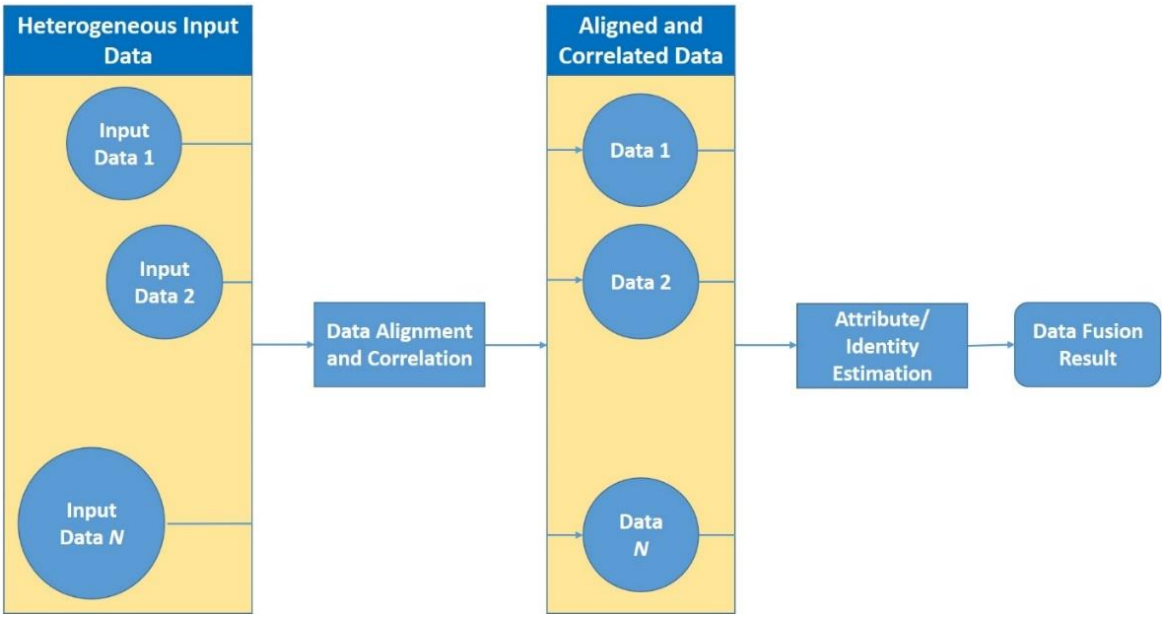


Figure 2.9: The generalised data fusion process (Adapted from Schmitt & Zhu, 2016, p. 9)

In general, remote sensing data/image fusion techniques are classified into three groups according to the level of processing involved: pixel/data level, feature level and decision level as shown in Figure 2.10 (Ghassemian, 2016; Mura et al., 2015; Pohl & Van Genderen, 1998; Yang et al., 2007; Zhang, 2010a).

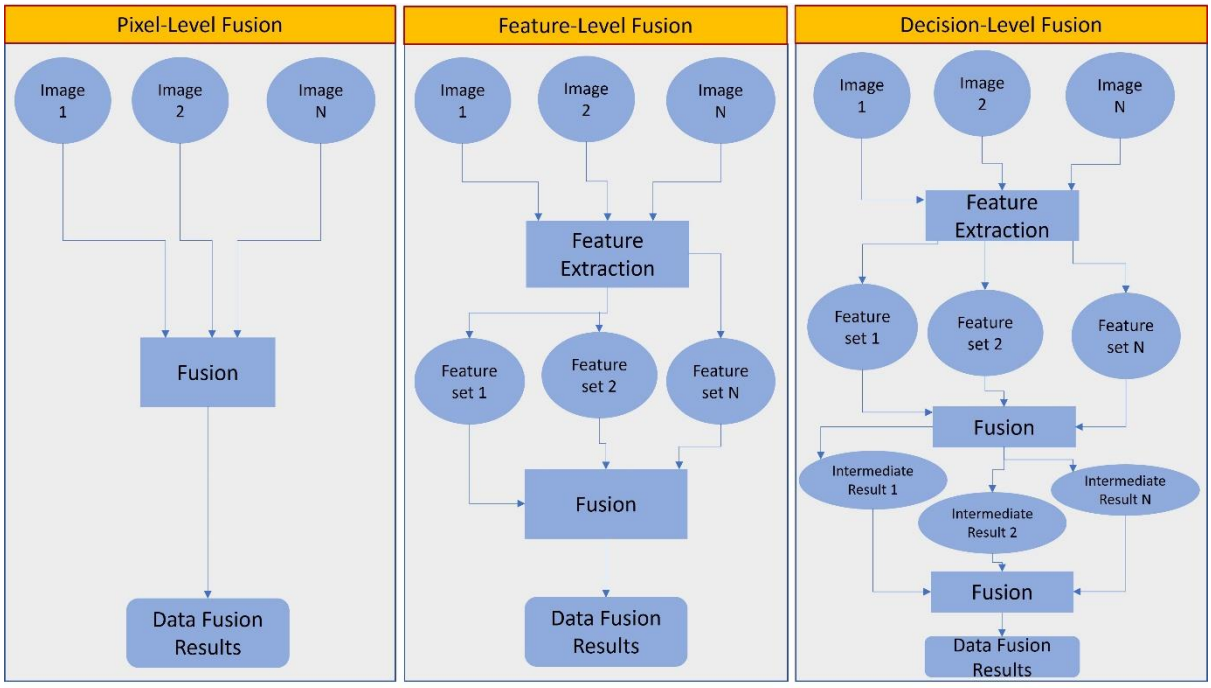


Figure 2.10: The three levels of image fusion (Adapted from Pohl & Van Genderen, 1998, p. 825; Schmitt & Zhu, 2016, p. 11)

2.6.1 Pixel level fusion

Pixel level (or raw data level) fusion is the most basic processing level of fusion for merging or combining measured physical parameters (Pohl & Van Genderen, 1998). Generally, it is designed to merge or combine multiple input images into a fused image, which is expected to contain more information, beyond what is provided by the input images (Li et al. 2017). Apart from the case of pansharpening of optical images, pixel level fusion can also be applied to multi-temporal data for highlighting spatio-temporal changes (Zhang, 2010). Yang et al. (2007) advise that the selection of a suitable pixel-level image fusion technique/algorithm depends on its merits, its relevance to applied situations and the source data characteristics. Three conditions must be satisfied in the fusion process: elimination of noise and irrelevant information, preservation of relevant information, and minimisation of inconsistencies and artefacts in the fused image (Ghassemian, 2016; Rockinger, 1996). Pixel level image fusion has been widely used for DEM fusion (Okolie & Smit, 2022).

2.6.2 Feature level fusion

Feature level fusion is a higher level of fusion in which extracted features are combined into feature maps for further processing (Zhang, 2010). The features are related to the properties of the source images and are dependent on the environment and neighbourhood (Mangolini, 1994; Pohl & Van Genderen, 1998). In digital image processing, feature-level fusion usually requires a co-registration of the source images and the feature maps are subsequently subjected to further processing to derive the required information (Zhang, 2010). Examples of feature level fusion methods include: multirate filter banks (Ghassemian, 2001), softmax regression-based feature fusion (Bai et al., 2015), learning-based super-resolution fusion (Song et al., 2015) and retina based multi-resolution (Ghassemian, 2001), layer stacking and ensemble learning methods (Pohl & Van Genderen, 2016). Feature-level fusion has also been adopted for DEM enhancement (e.g. Bhardwaj et al., 2019; Girohi & Bhardwaj, 2022).

2.6.3 Decision level fusion

Decision level fusion is used to derive a fused decision through the integration of outputs from several algorithms (Zhang, 2010). Decision rules are applied to resolve conflicts in the obtained information and to harmonise the interpretation (Pohl & Van Genderen, 1998; Shen, 1990). When the outcomes from algorithms are expressed as decisions, it is known as hard fusion, but when the outcomes are given as scores (confidences), it is referred to as soft fusion (Zhang, 2010). Examples of fusion methods at decision level include: voting (Jimenez & Morales-Morell, 1999), hybrid methods based on consensus (Benediktsson, 1997), Bayesian inference

(Mahmoudi et al., 2015), fuzzy decision rule (Fauvel et al., 2006), and Dempster-Shafer (Zhu et al., 2002).

2.7 Machine Learning in Remote Sensing Data Fusion

Machine Learning (ML) is "a process in which computing systems learn from data and use algorithms to execute tasks without being explicitly programmed" (Wazid et al., 2022, p. 314). Specifically, ML is a process that lets machines or computer systems interact with the world like humans, and to gain experience and learn with little human intervention (Meng et al., 2020). The ML approach is a subdivision of Artificial Intelligence (AI) that deals with the design of techniques or algorithms to learn from machine-readable data (Lary et al., 2016). ML allows the computer to 'learn' using given data, but without the need to thoroughly and explicitly program every problem (Meng et al., 2020). ML and AI methods have successfully been applied in recent years for solving problems in many areas such as supervised and unsupervised image classification and image recognition. Other applications include retrieval algorithms, detection of diseases in crops, bias correction, code acceleration, rock type classification and classification of pelagic habitats (Lary et al., 2018). ML includes several models that can perform non-linear, multi-variate, non-parametric regression or classification (Lary et al., 2016). When combined with remote sensing, ML can help improve the understanding of environmental systems.

2.7.1 Machine learning-based data fusion for DEM enhancement

Several methods have been proposed and adopted for correcting the elevation bias in DEMs. A common approach for DEM correction is the use of linear regression (Su & Guo, 2014; Olajubu et al., 2021; Pakoksung & Takagi, 2015; Preety et al., 2022). Other strategies have been proposed (e.g., Audenino et al., 2001; Bagheri et al., 2017, 2018b; Bhardwaj et al., 2019; Fu et al., 2016; Okolie & Smit, 2022). Nowadays, supervised ML provides the means for modelling complex relationships between variables, and has been deployed by researchers in a variety of fields (Hancock & Khoshgoftaar, 2020; Kotsiantis, 2007).

Some recent publications have investigated the widespread application of machine learning for DEM enhancement. For example, Wendi et al. (2016) improved the accuracy of 90 m SRTM DEM using an artificial neural network (ANN) that modelled the relationship between SRTM DEM errors and satellite imagery spectral signatures representing various land cover classes. Kulp & Strauss (2018) went further by integrating additional variables such as neighbourhood elevation values, vegetation cover indices, slope and population density into an ANN system

for DEM correction (Figure 2.11). Kim et al. (2020) incorporated building footprints and multispectral imagery into an ANN system for improvement of SRTM DEM. These studies recorded modest achievements in terms of error reduction in the output DEMs.

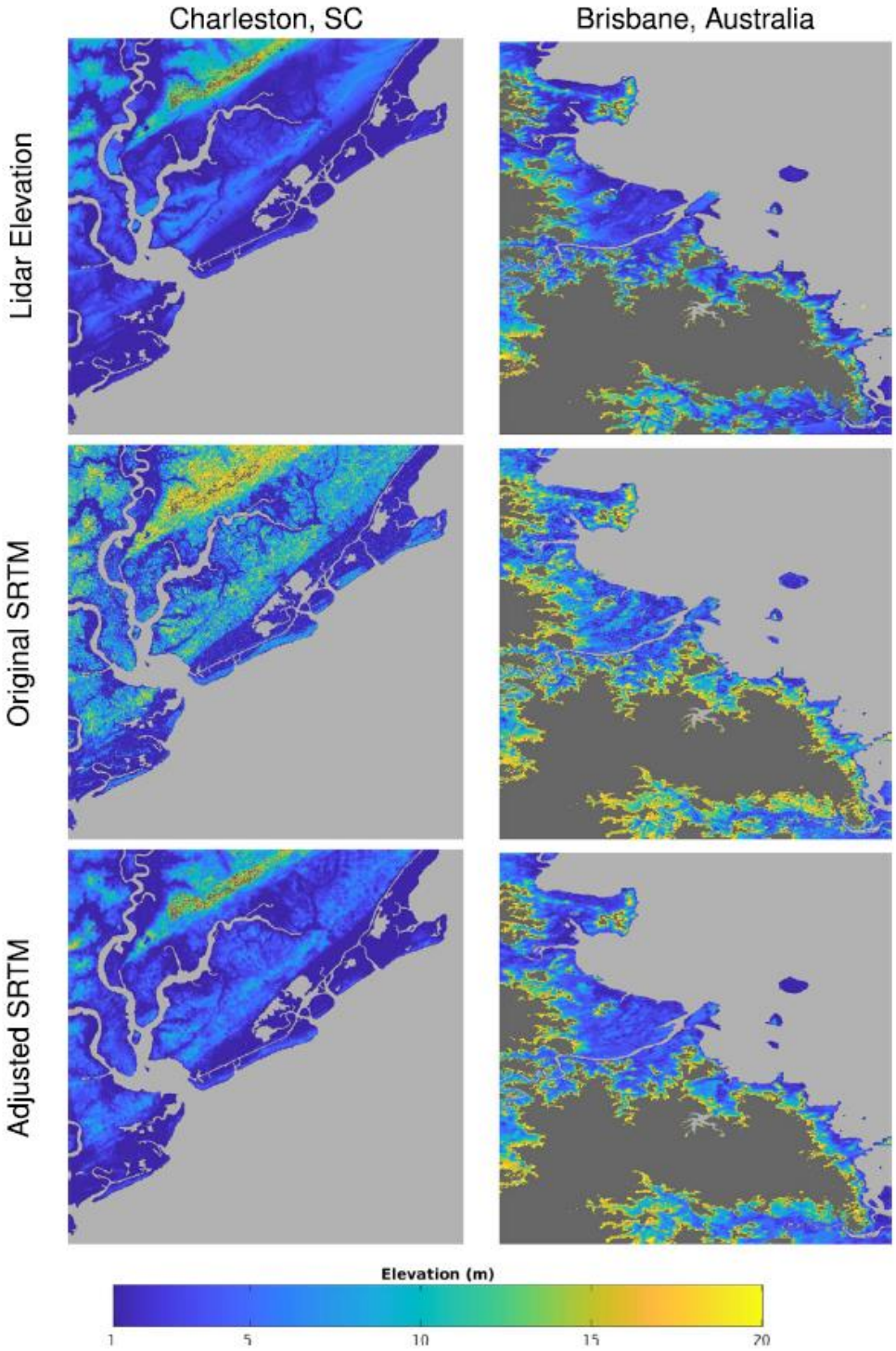


Figure 2.11: View of corrected SRTM DEM at two different sites in the USA and Australia (Source: Kulp & Strauss, 2018, p. 235)

Random forest regression has also been used in several instances, e.g., forest correction of SRTM (Su et al., 2015), urban correction of MERIT DEM (Liu et al., 2021), and SRTM DEM improvement for flood modelling (Meadows & Wilson, 2021). The proposed correction of MERIT DEM by Liu et al. (2021) reduced its vertical errors by 15 – 67% and 57 – 76% when using data from within and outside the city respectively. More recently, Hawker et al. (2022) used random forest in the generation of a globally corrected version of Copernicus DEM known as the Forests and Buildings removed Copernicus DEM (FABDEM; Figure 2.12). Hawker et al. (2022) posited that their method reduced the mean absolute error (MAE) in forests from 5.15 to 2.88 m and in built-up areas from 1.61 to 1.12 m.

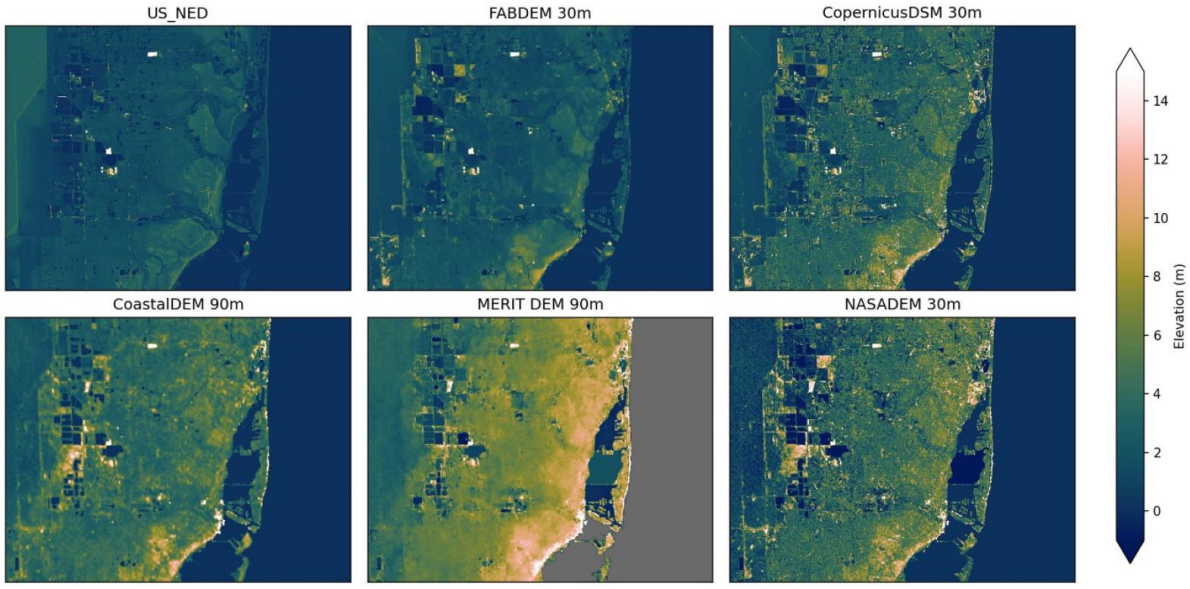


Figure 2.12: Comparison of FABDEM with different DEMs over a region of Florida, USA (Source: Hawker et al., 2022, p. 7)

2.7.2 Review of ML-based approaches for DEM enhancement

Table 2.1 shows a summary of studies that applied machine learning for DEM enhancement, brief descriptions of the approaches, and the summary of results. Usually, the input layers into the machine learning models are the elevation values from the DEM and values of other parameters that are known to influence or have relationships with the DEM error. While some researchers set the elevation/height error as the target layer (e.g., Kulp & Strauss, 2018; Liu et al., 2021), others utilized the groundtruth or reference DEM (RefDEM) as the target layer (e.g., Chen et al., 2020; Kasi et al., 2020; Kim et al., 2019). In another study by Salah (2021), probability images were adopted as the target layer. The recorded improvements vary depending on the studied landscape. The most common ML techniques used include random forest and neural networks.

Table 2.1: Summary of studies applying machine learning for DEM enhancement

Reference	Country of study	Brief description	Landscape	Datasets/layers	Summary of results
Kasi et al. (2020)	India	Combination of ML algorithms: genetic programming (GP) and artificial neural network (ANN), and interpolation techniques for DEM improvement. After training and testing, the predicted elevation values were interpolated into gridded DEMs using interpolation.	Urban, agricultural, water bodies, hillslope, grassland	**Datasets:** SRTM 30m, CARTOSAT 10m, CARTOSAT 30m, GCPs. **Input layers:** SRTM 30m elevations, CARTOSAT 10 m elevations, CARTOSAT 30m elevations. **Target layer:** GCP elevations	The root mean square error (RMSE) of the generated DEMs improved by 47% and 35%, compared to the SRTM and CARTOSAT DEMs respectively.
Chen et al. (2020)	China	Comparative assessment of multiple linear regression (MLR), generalized regression neural network (GRNN), random forest (RF), and back propagation neural network (BPNN) for correction of SRTM DEM.	Mixed land uses	**Datasets:** 90 m SRTM3 DEM v4.1, ICESat/GLAS elevation data, 30 m land use map. **Input layers**: SRTM elevations, slope, aspect, terrain relief, land use. **Target layer**: ICESat heights	In one of the experiments, RMSE changed from 10.05m in SRTM3 to 9.50m (MLR), 9.47m (BPNN), 10.38m (GRNN) and 9.21m (RF).

Ma et al. (2020)	China	Comparison of linear stepwise regression and BPNN for SRTM DEM improvement. The improvement system exploited the correlations between the normalised difference vegetation index (NDVI), potential solar radiation (PSR) and vertical drought index, surface material distribution and topography.	Cultivated land	**Datasets:** SRTM DEM, 5m DEM extracted from optical stereo-images (ZY-3) (for comparison), multispectral imagery, unmanned aerial vehicle (UAV)-derived DEM. **Input layers:** SRTM DEM elevations, slope, aspect, slope position, PSR, NDVI, PDI (perpendicular drought index). **Target layer:** UAV DEM elevations	RMSE reduced from 15.25m to 0.94m at the first site, and from 14.95m to 1.23m at the second site.
Salah (2021)	Egypt	Improvement of SRTM DEM accuracy using inverse probability weighted interpolation (IPWI), ANN and Sentinel-2 multispectral imagery.	Urban	**Datasets:** 30 m SRTM DEM v3, Sentinel-2 imagery, 1 m accuracy reference DEM. **Input layers:** SRTM DEM elevations, Sentinel-2 spectral bands 2, 3, 4 and 5, texture attributes from Sentinel-2 bands (contrast, dissimilarity, homogeneity, mean, entropy, range), SRTM DEM slope and aspect. **Target layer:** Probability images	RMSE reduction of ~29.35%.

Kim et al. (2019)	Malaysia, Singapore, Indonesia	An ANN framework for error reduction of SRTM DEM with the integration of Sentinel-2 multispectral imagery.	Tropical rainforest	**Datasets:** SRTM, multispectral imagery, 12 m TanDEM-X. **Input layers:** Sentinel-2 spectral bands, SRTM elevations. **Target layer:** TanDEM-X elevations	42.3% (7.8 m to 4.5 m) reduction in RMSE of SRTM.
Hu & Ji (2022)	China	Comparative analysis of BPNN, RF, gradient tree lifting and ensemble learning for improvement of STRM v1 (30 m), STRM v3 (90 m), AW3D30 (30 m), and ASTER GDEM (30 m).	Mountains, hills, plains and lakes.	**Datasets:** STRM v1 DEM (30 m), STRM v3 DEM (90 m), AW3D30 (ALOS World 3D-30 m) data, ASTER GDEM (30 m), ICESat/GLAS. **Input layers:** ICESat/GLAS: latitude and longitude, topography, aspect, slope, land use. **Target layer:** ICESat/GLAS – DEM elevation difference	**GBT model:** SRTM1 RMSE reduced by 16.43 % from 10.12 to 8.45 m. SRTM3 RMSE reduced by 6.7% from 14.65 to 13.67 m. **RF model:** AW3D RMSE reduced by 14.67 % from 9.05 to 7.72 m. ASTER RMSE reduced by 10.58 % from 11.19 to 10.00 m.

Wendi et al. (2016)	Singapore	Improvement of SRTM 90 m DEM using a standard feed forward neural network and Landsat 8 multispectral imagery.	Forested areas	**Datasets:** SRTM 90 m, Landsat 30 m imagery, 1 m accuracy DEM. **Input layers:** Original SRTM elevations Landsat-8 11 bands. **Target layer:** RefDEM heights	~68% (from 13.9 m to 4.4 m) and ~52% (from 14.2 m to 6.7 m) RMSE reduction at two different sites.
Hawker et al. (2022)	Global	Application of random forest for the reduction of forest and building bias in the 30 m Copernicus DEM.	Built-up areas, forests	**Datasets:** Forest heights, forest height image filters, land cover 100m – tree cover, ICESat-2 canopy heights, travel times, night lights, WorldPop constrained, GHS urban centre databases, World settlement footprint, Copernicus DEM image filters	Reduction of mean absolute error (MAE) in forests from 5.15 to 2.88 m, and in built-up areas from 1.61 to 1.12 m.
Girohi & Bhardwaj (2022)	India	A feed forward neural network for fusion of InSAR-based DEMs.	Diverse land cover	**Datasets:** Sentinel 1A/1B, Sentinel 2A, TanDEM-X 90 m DEM, toposheets, ICESat-2 ATL08 photon data. **Input layers:** DEM elevations, slope, aspect, TPI, TRI, VRM.	Significant improvement in RMSE (an improvement factor of 82.62% in hilly areas and 94.65% in plain areas).

Kim et al. (2021)	France, Singapore, Vietnam.	An ANN DEM enhancement scheme for improving the accuracy of 12 m TanDEM-X.	Built-up area, vegetation, and water bodies	**Datasets:** TanDEM-X DEM (12 m), Sentinel 2 multispectral imagery (10 – 60 m resolution), 1 m DEM (1 m spatial resolution and 40 cm vertical accuracy). **Input layers:** TanDEM-X elevations, NDVI, normalised difference built-up index (NDBI), normalised difference water index (NDWI), OpenStreetMap (OSM) building footprints. **Target layer:** Ground truth	21% to 65% RMSE reduction.
Kim et al. (2020)	France, Singapore	An ANN DEM enhancement scheme for improving the accuracy of 30m SRTM.	Urban	**Datasets:** SRTM 30 m, Sentinel 2 multispectral imagery band 2 – band 8A, 1m DEM (1 m spatial resolution and 40 cm vertical accuracy). **Input layers:** SRTM elevations, spectral reflectance values of Sentinel 2 multispectral imagery, OpenStreetMap (OSM) building footprints **Target layer:** Ground truth	~38% RMSE reduction.

Liu et al. (2021)	China, Germany, United Kingdom	Random forest was used for the treatment of building offsets/errors in the MERIT DEM.	Urban	**Datasets:** MERIT DEM, OpenStreetMap buildings, night-time lights, WorldPop 2015, reference DTM. **Input layers:** Satellite night-time lights, global population density, building density, building height, slope, neighbouring elevation values **Target layer:** Height error	Error reduction by 15–67%, and 57 – 76% respectively, in two separate experiments.
Kulp & Strauss (2018)	USA, Australia	A multilayer perceptron ANN was used to carry out a 23-dimensional vertical error regression analyses for DEM improvement.	Coastal areas	**Datasets:** 30 m SRTM 3.0 elevations, 90 m SRTM 2.1 elevations, Landscan 2010, Canopy height, MODIS VCF, ICESat elevations, NOAA Coastal LiDAR, Australia digital LiDAR, SRTM Water Body Dataset (SWBD). **Input layers:** SRTM elevations, slope, population density, tree canopy height, vegetation density. **Target layer:** SRTM-ICESat elevation error	Mean vertical bias reduced from 3.67 m to less than 0.01 m (coastal US), and from 2.49 m to 0.11 m (Australia). ∼50% RMSE reduction from 5.36 m to 2.39 m (US), and from 4.15 m to 2.46 m (Australia). Global bias reduced from 1.88 m to −0.29 m, and RMSE from 4.28 m to 3.08 m.

2.8 Tree-based Ensemble Machine Learning Algorithms

A decision tree is "a hierarchical splitting mechanism for classifying objects, and it provides a straightforward way to interpret and understand relationships between objects at different levels of detail" (Miao et al., 2011, p. 1829). Due to their ease of interpretation (Kotsiantis, 2013), several decision makers tend to feel more comfortable working with decision trees. Commonly used decision trees are ID3 (Chefn et al., 2009; Hssina et al., 2014), C4.5 (Quinlan et al., 1994) and the classification and regression tree (CART) (Loh, 2011; Speybroeck, 2012). Tree-based ensemble algorithms are generally divided into bagging and boosting. Bagging (known as bootstrap aggregating) is "a method for generating multiple versions of a predictor and using these to get an aggregated predictor" (Breiman, 1996, p. 123). By this approach, bagging leads to more stable results with decreased variance (Geurts et al., 2006; Zhang et al., 2022). Boosting algorithms "use a forward stagewise process to transform weak learners into strong learners by increasing the weights of training samples that were mistakenly identified or wrongly calculated in a successive iteration" (Zhang et al., 2022, p. 3). The final output is a combination of the results from multiple iterations (Elith et al., 2008; Zhang et al., 2022). Boosted trees are described as "the most effective off-the-shelf nonlinear learning method for a wide range of application problems" (Johnson & Zhang, 2014 , p. 942). Tree-based ensemble models can provide better accuracy than neural networks for many applications, and have outperformed deep learning in several experiments involving tabular data (Lundberg et al., 2020).

Gradient Boosted Decision Trees (GBDTs) are very powerful for addressing machine learning problems and have achieved impressive results in so many applications. Gradient boosting is excellent for handling heterogeneous, complex and noisy datasets, and for weather forecasting, recommendation systems, web search etc. (Caruana & Niculescu-Mizil, 2006; Roe et al., 2005; Zhang & Haghani, 2015; Prokhorenkova et al., 2018). Essentially, gradient boosting is "a process of constructing an ensemble predictor by performing gradient descent in a functional space" (Prokhorenkova et al., 2018, p. 1). The theoretical underpinnings for gradient boosting are well documented (e.g., Prokhorenkova et al., 2018).

Recent studies have indicated that GBDTs can outperform deep learning when applied to tabular datasets (e.g., Borisov et al., 2022; Kadra et al., 2021; Shwartz-Ziv & Armon, 2022). For example, in a soybean trait prediction study by Gill et al. (2022), the prediction performance of XGBoost or random forest outperformed deep learning models in 93% of

prediction sets. Also, Rewicki et al. (2023) showed that classical machine learning methods generally outperformed deep learning techniques across a range of anomaly types. The following sections present a review of ten tree-based ensemble ML algorithms.

2.8.1 Random forest

Random forest (RF) is a versatile and general-purpose supervised learning algorithm. It "combines several randomized decision trees and aggregates their predictions by averaging" (Biau & Scornet, 2016, p. 197). Breiman (2001) upgraded the previous bagging tree method to the RF algorithm. Through a random feature selection, the correlation between trees is reduced, thus leading to better predictions (Miao et al., 2011). RF operates on a "divide and conquer" principle, and can be adapted to various learning tasks and large-scale problems, and has displayed excellent performance in numerous settings (Biau & Scornet, 2016). The theory behind the RF algorithm is presented in Breiman (2001) and Biau & Scornet (2016), and its use in remote sensing has been reviewed by Belgiu & Drăgu (2016). Random forests have been successfully applied in various remote sensing practical problems, e.g., mapping of glacial lakes (Wangchuk & Bolch, 2020), geological mapping (Albert & Ammar, 2021), geoscience data analysis and modelling (Talebi et al., 2022), population estimation (Georganos et al., 2019), wetland classification (O'Neil et al., 2019), automatic crater detection (Wang et al., 2021), and DEM classification (Thong et al., 2012). Random forest has also been adopted by several researchers for DEM correction, e.g. forest correction of SRTM (Su et al., 2015), urban correction of MERIT DEM (Liu et al., 2021), SRTM DEM improvement for flood modelling (Meadows & Wilson, 2021), and forests and buildings removal from Copernicus DEM (Hawker et al., 2022).

2.8.2 Bagging meta-estimator

A bagging regressor (bagging meta-estimator) is "an ensemble meta-estimator that fits base regressors each on random subsets of the original dataset and then aggregate their individual predictions (either by voting or by averaging) to form a final prediction" (Scikit-Learn, 2023b, para. 1). The bagging meta-estimator encompasses several published works (Breiman, 1996, 1999; Ho, 1998; Louppe & Geurts, 2012). In many instances, bagging can improve a single model, without adapting the underlying base algorithm (Scikit-Learn, 2023a).

2.8.3 Adaptive boosting

For years, the AdaBoost (adaptive boosting) algorithm was a popular boosting algorithm (Schapire, 2003; Barrow & Crone, 2016). The core idea behind AdaBoost is the combination

of weak classifiers with different weights into a strong classifier (Liu et al., 2019). A series of individual classifiers is iteratively produced and each classifier in the ensemble attempts to classify the training data accurately (Kadavi et al., 2018). A weight adjustment policy is used at each round (Wu & Nagahashi, 2015). The final model is derived from the weighted sum of all classifier base models. One of the merits of AdaBoost is its resistance to overfitting but on the downside, it does not handle noise well (Bauer & Kohavi, 1999; Miao et al., 2011). AdaBoost can be used for evaluating the relevance of variables by examining how often they are selected by weak learners (Kadavi et al., 2018).

AdaBoost and its variants have been used in diverse fields, e.g., face detection (Viola & Jones, 2001), hand detection (Guo et al., 2012; Kölsch & Turk, 2004), and human detection (Xu et al., 2012). Some demonstrated applications of AdaBoost in remote sensing include: aerial LiDAR data classification (Lodha et al., 2007), classification of ecological zones (Miao et al., 2011), terrain classification (Nguyen & Woo, 2014), landslide susceptibility mapping (Kadavi et al., 2018; Wu et al., 2020), and classification of power facility point clouds (Liu et al., 2019).

2.8.4 Gradient boosting machine

The capability of gradient boosted regression trees for producing robust, competitive and interpretable procedures for regression is well known (Friedman, 2001). The gradient boosting regressor "builds an additive model in a forward stage-wise fashion; it allows for the optimisation of arbitrary differentiable loss functions. In each stage a regression tree is fit on the negative gradient of the given loss function" (Scikit-Learn, 2023c, para. 1). The gradient boosting regressor implementation in Scikit-Learn supports several loss functions for regression.

2.8.5 Extreme gradient boosting

Extreme gradient boosting (XGBoost) is a "scalable end-to-end tree boosting system" that has achieved wide recognition and provides competitive results in several machine learning challenges (Chen & Guestrin, 2016; p. 785). XGBoost is "an ensemble of decision trees; it sequentially builds short and simple decision trees as each tree tries to improve the performance of the previous one" (Safaei et al., 2022, p. 9). The scalability of XGBoost is attributed to several innovations such as (Chen & Guestrin, 2016): (i) an innovative tree learning algorithm; (ii) a theoretically justified weighted quantile sketch procedure; (ii) parallel and distributed computing, and (iv) out-of-core computation. Several studies have shown the potential of XGBoost for remote sensing tasks, e.g., estimating forest growing stock (Huang et al., 2022),

patch-based image classification (Ozturk & Colkesen, 2021), and soil moisture mapping (Ågren et al., 2021).

2.8.6 Light gradient boosting machine

To address the efficiency and scalability drawbacks of previous engineering optimisations used in GBDTs, Microsoft (Ke et al., 2017) introduced an improved GBDT framework, the light gradient boosting machine (LightGBM). LightGBM is "mainly featured by the decision tree algorithm based on gradient-based one-side sampling (GOSS), exclusive feature bundling (EFB), and a histogram and leaf-wise growth strategy with a depth limit" (Li et al., 2022, p. 3). Summarily, GOSS and EFB increase the speed of training (Safaei et al., 2022). The advantages of LightGBM include (McCarty et al., 2020): (i) faster speed of training (ii) highly efficient (iii) lower memory usage (iv) improved accuracy (v) better compatibility with large datasets, and (vii) support for parallel, distributed and GPU learning. A detailed theoretical background of LightGBM is presented in Ke et al. (2017).

2.8.7 Histogram-based gradient boosting

Inspired by LightGBM, a histogram-based gradient boosting regression tree was introduced in Scikit-Learn 0.21. This algorithm is reputed to be faster than the gradient boosting regressor for large datasets, and has built-in support for missing values (Scikit-Learn, 2023a). The histogram-based method "buckets continuous feature values into discrete bins and uses these bins to construct feature histograms during training" (Padhi et al., 2021, p. 12). Moreover, the algorithm increases the speed of training but without loss of accuracy (Padhi et al., 2021). An implementation of histogram-based gradient boosting is provided in the Scikit-Learn library. Applications include accident severity prediction (Kashifi & Ahmad, 2022) and financial market forecasts (Padhi et al., 2021).

2.8.8 Categorial boosting

Categorical Boosting (CatBoost) debuted in 2018, and is well-suited for handling categorical data (Hancock & Khoshgoftaar, 2020). Two important algorithmic advances were introduced in CatBoost: ordered boosting, and the processing of categorical features using a novel algorithm (Prokhorenkova et al., 2018). CatBoost employs ordered boosting along with ordered target statistics that are based on the ordering principle to address target leakage and the problems of prediction shift which are caused by differences in the distribution of train and test datasets (Wang et al., 2022). CatBoost is implemented as an open-source library, and has outperformed XGBoost and LightGBM in several machine learning tasks and competitions

(Prokhorenkova et al., 2018; Wang et al., 2022). CatBoost can provide impressive results even with limited data and resources (Safaei et al., 2022). The key algorithmic techniques behind CatBoost are presented in the seminal work by Prokhorenkova et al. (2018).

Since its release in 2018, CatBoost has found wide use in applications such as spatial data quality assessment (Takacs, 2018), water depth inversion (Ruiyao et al., 2022), abrasive belt grinding (Wang et al., 2022), astronomy (Khramtsov et al., 2019), finance (Xia et al., 2020), medicine (Yang & Bath, 2020), biology (Kolesnikov et al., 2019), electrical utilities fraud (Punmiya & Choe, 2019), meteorology (Fan et al., 2020), psychology (Abolfazli et al., 2020), traffic engineering (Liu et al., 2020), cyber-security (Bakhareva et al., 2019), bio-chemistry (Yi et al., 2019), marketing (Koehn et al., 2020) and electricity theft detection (Hussain et al., 2021). Applications in remote sensing include: particulate matter estimation (Ding et al., 2021), water quality monitoring (Chen et al., 2023). estimating forest growing stock (Huang et al., 2022), hyperspectral image classification (Samat et al., 2021), patch-based image classification (Ozturk & Colkesen, 2021), and sea surface pressure estimation (Sun et al., 2021).

2.8.9 Natural gradient boosting

Natural Gradient Boosting (NGBoost) is a gradient boosting algorithm developed for generic probabilistic prediction. NGBoost "generalizes gradient boosting to probabilistic regression by treating the parameters of the conditional distribution as targets for a multiparameter boosting algorithm" (Duan et al., 2020, p. 1). The working principle of NGBoost is well documented (e.g., Duan et al., 2020). It is applicable to multiple tasks such as classification, regression and survival prediction (Kavzoglu & Teke, 2022). NGBoost has been implemented as a Python library that is built on top of Scikit-Learn (Ngboost, 2023). The algorithm is gaining traction within the machine learning community and recent applications include: landslide susceptibility mapping (Kavzoglu & Teke, 2022), solar power forecasting (Mitrentsis & Lens, 2022), infrastructure performance prediction (Chen et al., 2022), and runoff probability prediction (Shen et al., 2022). It is reported that NGBoost outperforms existing methods for probabilistic prediction (Duan et al., 2020).

2.8.10 Regularised greedy forest

The Regularized Greedy Forest (RGF) is a method that "directly learns decision forests via fully-corrective regularized greedy search using the underlying forest structure" (Johnson & Zhang, 2014, p. 942). RGF integrates tree-structured regularization and a fully-corrective

regularized greedy algorithm (Johnson & Zhang, 2014). It has been shown to outperform GBDTs in several machine learning challenges (Joseph, 2020).

2.9 Explainability of Machine Learning Algorithms

Explainable Artificial Intelligence (XAI) is concerned with methods that explain and interpret the performance of ML algorithms. Explainability is associated with the internal mechanics and logic that are within a machine learning system while interpretability deals with the intuition behind the model outputs (Adadi & Berrada 2018; Linardatos et al. 2020). Machine learning-based feature importance techniques are used to explain the importance or influence of the input variables, and provide an understanding of the predictions made by complex models (Nordin et al., 2023). Furthermore, partial dependence plots (PDPs) and shapley additive explanations (SHAP) plots are used for showing the relationship between the features and the predictions (e.g., Friedman, 2001; Nohara et al., 2015, 2022). The explainability of machine learning models has received so much wide attention in recent years that XAI has become a domain in the AI community (Nauta et al., 2023; Saleem et al., 2022). Extensive reviews of machine learning explainability and interpretability are available in the literature (Adadi & Berrada, 2018; Dosilovic et al., 2018; Linardatos et al., 2020; Javed et al., 2023; Saeed & Omlin, 2023; Schwalbe et al., 2023).

2.10 Chapter Summary

The importance of global DEMs for a myriad of environmental applications is evident from the literature review. Quality limitations of the DEMs such as systematic and random error were discussed. Several terrain parameters which reveal statistically significant differences on the terrain were also discussed. The advantage of machine learning for remote sensing data fusion with specific applications to DEM enhancement were reviewed. From available evidence, the more recent implementations of gradient boosting are yet to be tested for DEM enhancement. The literature review also buttresses the fact that well optimised machine learning solutions can outperform deep learning solutions. The body of knowledge and state-of-the-art in ML-based DEM enhancement is growing and could emerge as a unique field in the nearest future.

CHAPTER 3

THEORETICAL FRAMEWORK

3.1 Introduction

In this chapter, the established theories that support the techniques, tools and approaches adopted in this research and that are specifically relevant to the research questions are discussed. These include the theoretical underpinnings behind decision trees, bagging and boosting ensembles, gradient boosting, model optimisation and explainability. Collectively, the theoretical framework provides a roadmap for the formulated workflows in the methodology.

3.2 Theoretical Framework of Decision Trees

Decision trees are classified under supervised machine learning. They are "sequential models, which logically combine a sequence of simple tests; each test compares a numeric attribute against a threshold value or a nominal attribute against a set of possible values" (Kotsiantis, 2013, p. 261). Decision trees have been employed multifariously in numerous fields as prediction models. Anuradha & Gupta (2014, p. 2) describe the traditional structure of a decision tree: "the decision tree is a directed tree that comprise of a root node, internal node and leaf node. Root node has no incoming edges and leaf node has no outgoing edges. The internal node represents test on an attribute and branch represents the outcome of that test." The 'decisions' are contained within the leaf nodes (Williams, 2011). Examples of decision tree algorithms include: ID3 (Quinlan, 1986), CHAID (Rokach & Maimon 2005; Somvanshi et al., 2017) and CART (Breiman et al., 2017). Figure 3.1 shows the structure of a simple decision tree.

3.3 Tree-based Ensemble Algorithms

The concept of Ensemble learning (EL) is based on the creation of multiple models and combination of their results for a final output (Aksoy & Genc, 2023). This contrasts with the use of a single strong predictive model which is the most frequent approach in data-driven modelling (Natekin & Knoll, 2013). Numerous empirical and theoretical studies have established the advantages of combining EL techniques such as bagging or boosting, which often results in improved generalisation performance and several theoretical explanations have been brought forward (Breiman, 1996; Hastie et al., 2009; Kuncheva, 2004; Graczyk et al., 2010; Freund & Schapire 1997).

Ensemble methods are considered as state-of-the-art and have been used for solving numerous challenges in machine learning and remote sensing. Bagging and boosting are suitable for classification and regression problems. It has been posited that boosting algorithms often provide better predictive power than bagging algorithms. However, this assertion should be treated with caution since performance could vary on individual datasets (Bühlmann, 2012).

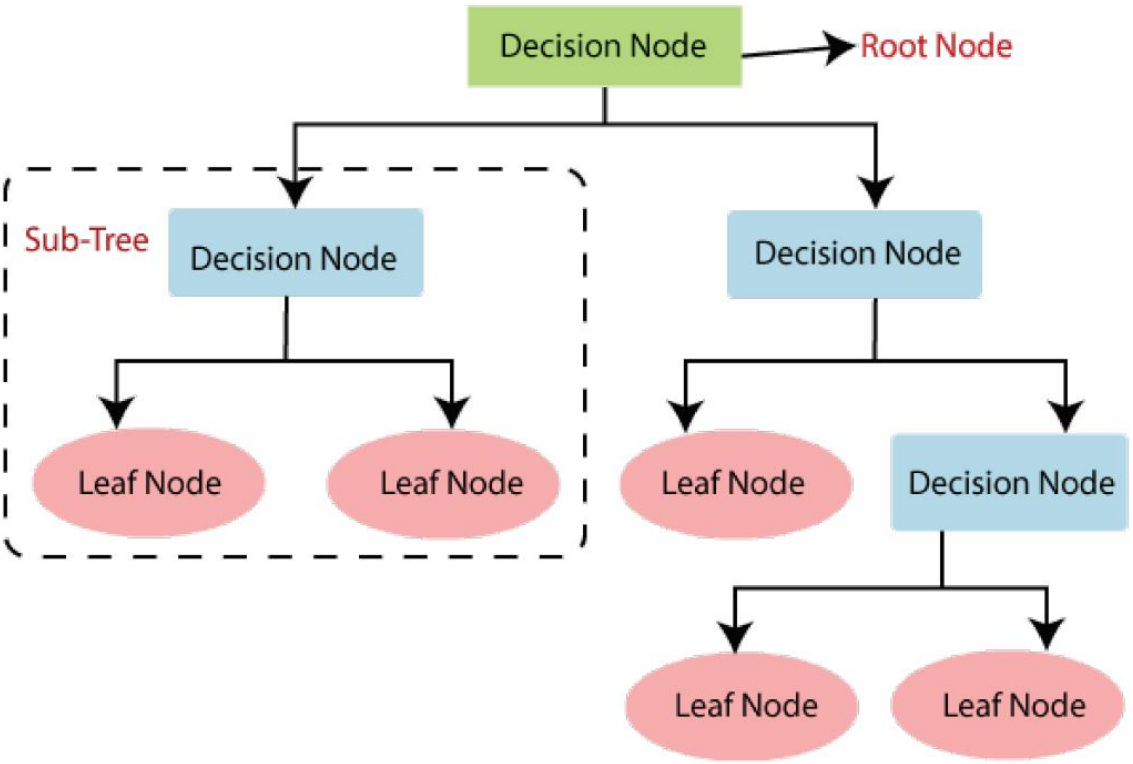


Figure 3.1: The structure of a decision tree (Source: Hafeez et al., 2021, p. 4)

3.3.1 Bagging

Bagging (known as bootstrap aggregating) is "a method for generating multiple versions of a predictor and using these to get an aggregated predictor" (Breiman, 1996, p. 123). It is intuitive, simple to implement and has good performance (Breiman, 1996; Polikar, 2006). To achieve diversity in bagging, bootstrapped replicas are used, i.e., "different training data subsets are randomly drawn—with replacement—from the entire training data. Each training data subset is used to train a different classifier of the same type. Individual classifiers are then combined by taking a majority vote of their decisions. For any given instance, the class chosen by most classifiers is the ensemble decision" (Polikar, 2006, p. 27). Figure 3.2 presents the pseudocode for bagging.

Algorithm: Bagging

Input:

- Training data S with correct labels $\omega_i \in \Omega = \{\omega_1, \ldots, \omega_C\}$ representing C classes
- Weak learning algorithm **WeakLearn**,
- Integer T specifying number of iterations.
- Percent (or fraction) F to create bootstrapped training data

Do $t = 1, \ldots, T$

1. Take a bootstrapped replica S_t by randomly drawing F percent of S.
2. Call **WeakLearn** with S_t and receive the hypothesis (classifier) h_t.
3. Add h_t to the ensemble, E.

End

Test: Simple Majority Voting – Given unlabeled instance **x**

1. Evaluate the ensemble $\mathrm{E} = \{h_1, \ldots, h_T\}$ on **x**.
2. Let $v_{t,j} = \begin{cases} 1, & \text{if } h_t \text{ picks class } \omega_j \\ 0, & \text{otherwise} \end{cases}$

   be the vote given to class ω_j by classifier h_t.
3. Obtain total vote received by each class

$$V_j = \sum_{t=1}^{T} v_{t,j}, \quad j = 1, \ldots, C$$

4. Choose the class that receives the highest total vote as the final classification.

Figure 3.2: The Bagging algorithm (Source: Polikar, 2006, p. 28)

3.3.2 Boosting

Schapire (1990) introduced an algorithm called boosting, for improving the performance of weak learners. The boosting approach has a sequential working structure (Aksoy & Genc, 2023). Boosting "creates an ensemble of classifiers by resampling the data, which are then combined by majority voting" (Polikar, 2006, p. 28). However, the resampling is done in such a way that it provides very informative training data for the subsequent classifier (Polikar, 2006). The boosting algorithm is shown in Figure 3.3. Unlike earlier boosting methods which were purely algorithm-driven, gradient boosting machines have established a connection with the statistical framework thus offering better explainability (Friedman, 2001; Friedman et al., 2000; Friedman, 2001; Freund & Schapire, 1997; Natekin & Knoll, 2013).

Algorithm: Boosting

Input:

- Training data S of size N with correct labels $\omega_i \in \Omega = \{\omega_1, \omega_2\}$;
- Weak learning algorithm **WeakLearn**.

Training

1. Select $N_1 < N$ patterns without replacement from S to create data subset S_1.
2. Call **WeakLearn** and train with S_1 to create classifier C_1.
3. Create dataset S_2 as the most informative dataset, given C_1, such that half of S_2 is correctly classified by C_2, and the other half is misclassified. To do so:
   a. Flip a fair coin. If Head, select samples from S, and present them to C_1 until the first instance is misclassified. Add this instance to S_2.
   b. If Tail, select samples from S, and present them to C_1 until the first one is correctly classified. Add this instance to S_2.
   c. Continue flipping coins until no more patterns can be added to S_2.
4. Train the second classifier C_2 with S_2.
5. Create S_3 by selecting those instances for which C_1 and C_2 disagree. Train the third classifier C_3 with S_3.

Test – Given a test instance **x**

1. Classify **x** by C_1 and C_2. If they agree on the class, this class is the final classification.
2. If they disagree, choose the class predicted by C_3 as the final classification.

Figure 3.3: The Boosting algorithm (Source: Polikar, 2006, p. 29)

3.4 Gradient Boosting Machines

Gradient boosting is an advanced concept of boosting. Gradient boosting machines (GBMs) construct new base-learners in such a way that they are very highly associated with the negative gradient of the loss function (Natekin & Knoll, 2013). The working principle of GBM-based regression is presented in Aksoy & Genc (2023). Specialised algorithms based on GBM include adaptive boosting (AdaBoost), light boosting machine (LightGBM), categorical boosting (CatBoost), histogram-based gradient boosting etc. GBM-based algorithms are based on the tree structure used in the decision tree (DT) algorithm (Aksoy & Genc, 2023). Below, the theoretical underpinnings of the three most popular techniques (XGBoost, LightGBM and CatBoost) are presented.

3.4.1 Extreme gradient boosting

The tree structures combined in the XGBoost algorithm are based on the Classification and Regression Tree (CART). In the case of regression, the CART methodology consists of three stages: (i) a regression tree is grown which overfits the data, (ii) a sequence of sub-trees is pruned from the overfitting tree, and (iii) a subtree which estimates the true regression function is selected from the sequence of subtrees (Klemelä et al., 2000). In XGBoost, the decision trees are built simultaneously in a multi-level fashion such that the tree attributes are sorted once in each stage (Omer & Shareef, 2022). The prediction scores of individual trees in the ensemble model are combined to derive the final score. The XGBoost framework adopts random sampling for improving the predictive power and for reduction of model variance (Jabeur et al., 2021):

$$Z = F(x_i) = \sum_{t=1}^{T} f_t(x_i) \quad (3.1)$$

Where,

(x_i) refers to the explanatory variables

$f_t(x_i)$ is the output function of each tree

A combination of parameters can improve the predictive power, and a regularisation term is added for model optimisation (Jabeur et al., 2021).

3.4.2 Light boosting machine

LightGBM adopts the gradient based oneside sampling (GOSS) strategy for splitting internal nodes, and exclusive feature bundling (EFB) for reducing input feature dimensions (Cui et al., 2021; Liu et al., 2023). Suppose a dataset has n instances, $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$,

Where,

$\{x_1, x_2, \dots, x_n\}$ represent the independent variables

$\{y_1, y_2, \dots, y_n\}$ represent the dependent variables

The predictive value, f(x) of the gradient boosted decision tree is derived from the sum of the outputs of a group of decision trees $h_t(x)$ (Liu et al., 2023):

$$f(x) = \sum_{t=1}^{T} h_t(x) \quad (3.2)$$

Where,

T is the no. of decision trees.

The fitting model finds an approximate function $\hat{f}$, whose relationship with $L(y, f(x))$ is:

$$\hat{f} = \arg\min_{f} E_{y,S} L(y, f(x)) \quad (3.3)$$

3.4.3 Categorical boosting

CatBoost adopts a modification of gradient boosting referred to as ordered boosting, in order to overcome target leakage (Dorogush et al., 2018; Jabeur et al., 2021). Another of its advantages is the performance of random permutations for estimating leaf values to overcome overfitting (Dorogush et al., 2018; Jabeur et al., 2021). In CatBoost, the base predictors are binary decision trees. The output is given as (Dorogush et al. 2018; Prokhorenkova et al., 2018; Jabeur et al., 2021; Gao et al., 2022):

$$Z = H(x_i) = \sum_{j=1}^{J} c_j 1_{\{x \in R_j\}} \quad (3.4)$$

Where,

$H(x_i)$ is a tree function of the explanatory/input variables x_i

R_j is the disjoint area corresponding to the leaves of the tree.

3.5 Bayesian Optimisation for Hyperparameter Tuning

Bayesian optimisation (BO) has emerged as a powerful solution for constructing optimal programs that compute better solutions or that run faster (Hoos, 2010; Shahriari et al., 2016). It has impacted a wide range of areas including experimental design (Azimi et al. 2012), automatic machine learning (Bergstra et al., 2011), combinatorial optimisation (Wang et al., 2013), robotics (Martinez-Cantin et al., 2008), and information extraction (Wang et al., 2014). The performance of ML algorithms can be improved through Bayesian hyperparameter optimisation. BO "builds a probability model of the objective function to determine the optimal hyperparameters in an informed manner, reducing the number of times the objective function needs to be run by choosing only the most promising set of hyperparameters".(Habtemariam et al., 2023, p. 9).

In BO, the objective function is minimised within a finite domain (Habtemariam et al., 2023).

$$x^* = \arg\min_{x \in X} f(x) \quad (3.5)$$

Where,

f(x) is a score to be minimized

X is the domain of the hyperparameter values

x^* is a combination of hyperparameters that produces the lowest value of the score f(x)

3.6 ML Explainability with Shapley values

The Shapley value was proposed by Lundberg and Lee (2017) for determining the feature importance of input variables. It is a fair profit allocation among the predictor variables and is dependent on their relative contributions to the outcome (Nohara et al., 2022; Roth, 1988). Consider the predictive model f, the jth feature with i feature combinations, j index, target feature x, and data streams D with matrix X. The derivation of the Shapley value for x (ϕ^i_j) follows (Gebreyesus et al., 2023):

$$\emptyset^i_j = \hat{f}(X + j) - \hat{f}(X - j) \quad (3.6)$$

Where,

$\hat{f}(x_{+j})$ is the prediction for target x with a random number of feature values, including jth feature, and $\hat{f}(x_{-j})$ is the coalition without the jth feature.

Generally, the Shapley value of the jth feature that is target x is (Gebreyesus et al., 2023):

$$\emptyset_j(x) = \frac{1}{n}\sum_{i=1}^{n}(\emptyset^i_j) \quad (3.7)$$

Using the same approach, the importance of all the features is derived and the ranking is done based on the Shapley value. SHapley Additive exPlanations (SHAP) take advantage of Shapley values to construct additive explanatory models that enhance the interpretation of machine learning algorithms (Lu et al. 2023). SHAP is based on the cooperative game-theory and is model-agnostic (Anjum et al., 2022; Nordin et al., 2023; Palar et al., 2023; Lu et al., 2023). With SHAP, underlying patterns in the data and feature interactions can be detected (Gebreyesus et al., 2023).

3.7 Chapter Summary

This chapter treated the theory, structure and working principles of decision trees, and the state-of-the-art tree-based ensemble machine learning. The sequence of operations involved in

bagging and boosting ensembles which enable the realisation of more accurate predictions from weak learners was also discussed. Also, the theoretical basis was presented for Bayesian optimisation, a powerful solution for improving model performance.

CHAPTER 4

RESEARCH METHODOLOGY

4.1 Introduction

Figure 4.1 presents the overarching framework for the six stages in the methodology, while detailed workflows are presented in subsequent sections. Although several stages and/or sub-stages are interlinked, the following core operations in each stage are summarised below.

1. **Stage 1 (Data acquisition):** The acquisition and compilation of the global DEMs and ancillary datasets.
2. **Stage 2 (Data alignment):** The DEMs are aligned through mosaicking and datum harmonisation (in preparation for quality assessment). The selected DEMs are further aligned through co-registration and resampling (in preparation for tree-based ensemble feature-level fusion and enhancement).
3. **Stage 3 (Quality assessment):** The quantitative and qualitative assessments of the qualities of four global DEMs to determine their qualities in various landscapes. This assessment informs the selection of DEMs for enhancement.
4. **Stage 4 (Feature extraction):** Terrain parameters are extracted from the DEMs and integrated with land cover parameters. All parameters are aligned with the source DEMs following the procedure in stage 2.
5. **Stage 5 (Feature-level fusion):** Tree-based ensemble learning systems are configured for fusing the extracted features to learn the spatial patterns of DEM error in five diverse landscapes. The learning process incorporates Bayesian optimisation for improving the probability of obtaining optimal solutions in the error prediction. Subsequently, the predicted errors are applied for correcting the vertical error (elevation error) of the selected DEMs.

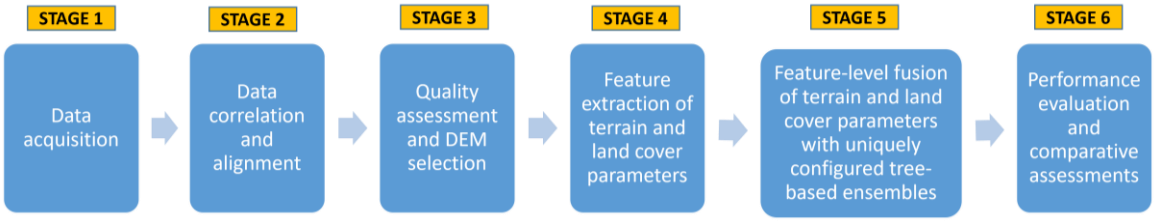


Figure 4.1: Generalised methodological framework

6. **Stage 6 (Performance evaluation):** A series of performance evaluations and comparative assessments are conducted to validate the proposed DEM enhancement scheme.

Several algebraic calculations were performed at different stages of the methodology implementation within ArcGIS software. For example, in the case of point data, elevation differencing was done using the field calculator, while raster operations were performed using map algebra.

4.2 Study Area

The study area is Cape Town, South Africa's most south-western city, with a land area of about 400 km^2 (Orimoloye et al. 2019). Cape Town (shown in Figure 4.2) is situated on the south-western coast of the Western Cape Province. The coastline is a mixture of sands, rocks and mountains. Topographically, Cape Town is exceptionally diverse. This sort of landscape presents diverse terrain contexts suited for the present study. The Western Cape is the fourth largest of the nine provinces in South Africa and is geographically located between longitudes $17°45'00'' - 24°15'00''$E and latitudes $30°20'00'' - 34°55'00''$S. It is bordered on the north by the Northern Cape and on the east by the Eastern Cape.

Five different landscapes in Cape Town were selected for the study (Figure 4.3). The sites are exceptionally diverse in terms of land cover and topography, providing a wide spectrum of geomorphological contexts well suited for the presented study.

4.2.1 Urban/industrial

The urban and industrial areas are located within the Cape Town metropolis, a large urban area with a high population density, industrial and business districts. The selected areas include the Parow and Elsies River Industrial areas, and several residential and commercial buildings.

4.2.2 Agricultural lands

These are agricultural fields with few settlements located along the floodplain of the Diep River. The Diep River, with its sources in the Riebeek-Kasteel Mountains, flows in a south-west direction towards Table Bay where it widens out forming a large wetland before entering the sea (Drakenstein Municipality, 2009). The Diep River is a sub-catchment within the Berg Catchment area, and is surrounded by informal settlements, industries, factories and farms (DWS, 2020; Gqomfa et al., 2022). The river catchment is low-lying and flat, and falls into the western lowland area of the Western Cape (DWS, 2002).

4.2.3 Mountain

The high flat-topped Table Mountain is part of the Cape Fold Belt; the mountain rises impressively from the surrounding ocean, and dominates the Cape Town landscape (Meadows & Compton, 2015).

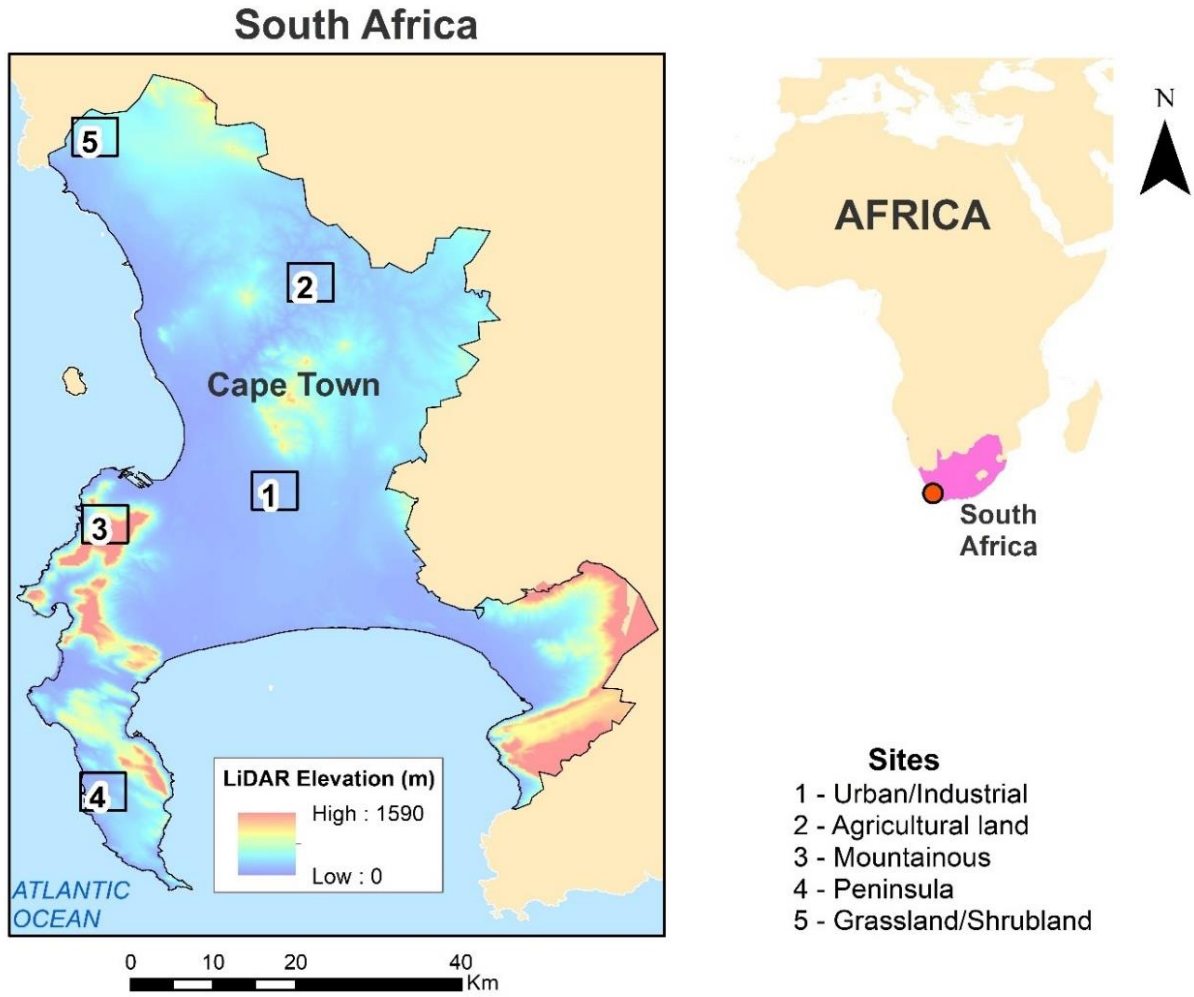


Figure 4.2: Maps showing the study area (Cape Town South Africa). The numbered boxes are the selected sites for the DEM quality assessment

4.2.4 Peninsula

The Cape Peninsula is topographically diverse, and forms part of the Cape Folded Belt (Cowling et al., 1966; Cowling et al., 1996; Simmons & Cowling, 1996). According to Cowling et al. (1996, p. 535), "the impressive Peninsula Mountain Chain is separated from the north-south trending Folded Belt on its eastern margin by the relatively warm waters of False Bay and the narrow sandy isthmus of the Cape Flats; on its western margin it plunges, sometimes precipitously, into the cold waters of the Atlantic Ocean."

4.2.5 Grassland/shrubland/dense bush

The general descriptions of these areas are summarized from the 2013 – 2014 South African National Land Land-Cover Dataset (DFFE, 2023) as follows:

a. **Grassland:** natural or semi-natural grass dominated areas, including open grassland, sparse bushland, transitional wooded grasslands and woodland areas.
b. **Low shrubland**: natural or semi-natural low shrubs, including degraded vegetation with significantly reduced vegetation cover.
c. **Dense bush/thicket**: natural or semi-natural areas dominated by trees and/or bushes, including dense bush, closed woodland, thicket, scrub forest, dense shrubs and mangrove swamps.

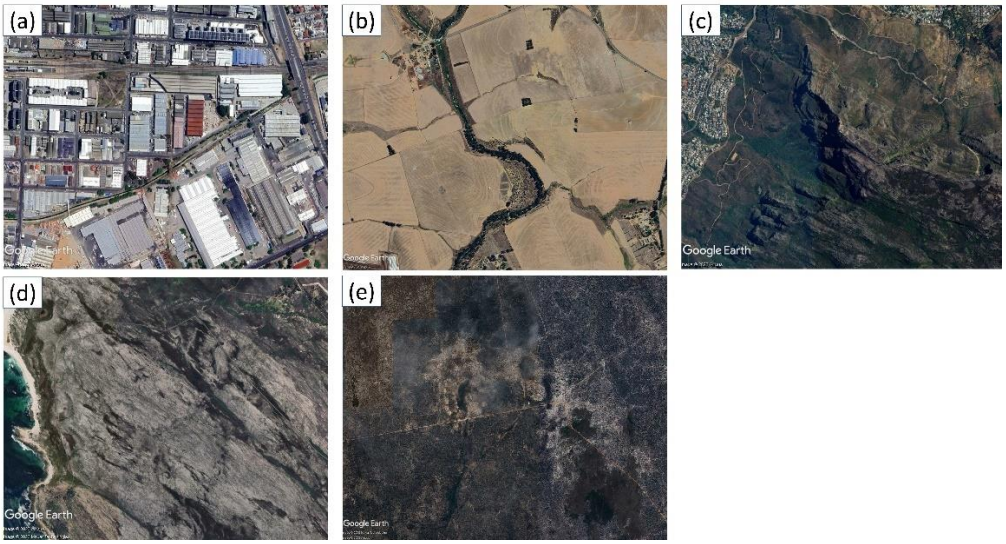


Figure 4.3: Some parts of the five landscapes considered in this research – (a) urban/industrial (b) agricultural (c) mountain (d) peninsula (e) grassland/shrubland. (Source: Google Earth)

4.3 Datasets, Terrain and Land Cover Parameters

The 30 m global DEMs are presented with their characteristics and tile identifiers (IDs) in Table 4.1 and Table 4.2 respectively. NASADEM and Copernicus GLO-30 were produced using Synthetic Aperture Radar interferometry (InSAR) while ASTER and AW3D are photogrammetric DEMs. The 2 m airborne LiDAR-derived DEM of Cape Town is adopted as the reference DEM. In the performance evaluation for lineament mapping, the 5 m airborne LiDAR-derived DEM is used for comparison.

4.3.1 NASADEM v001

The one arc-second resolution NASADEM is a modernisation of the Shuttle Radar Topography Mission (SRTM) data (Buckley et al., 2022). The original SRTM data was reprocessed with

enhanced algorithms, and more data was integrated from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and the Ice, Cloud, and Land Elevation Satellite – Geoscience Laser Altimeter System (ICESat GLAS) instruments. Significant reduction in voids were achieved in NASADEM through phase unwrapping improvements and the use of ICESat GLAS data for control (Buckley et al., 2020). Five different NASADEM products have been released to the global user community (Bettiol et al., 2021). NASADEM is distributed in 1° x 1° tiles and the coverage spans between 60°N and 56°S.

4.3.2 ASTER GDEM v3

The Advanced Space-borne Thermal Emission and Reflection Radiometer Global DEM (ASTER GDEM) is the result of a collaboration between the National Aeronautics and Space Administration (NASA) and the Ministry of Economy, Trade and Industry (METI) of Japan. ASTER GDEM was produced using imagery acquired from the NASA Terra mission. GDEM version 1 achieved an overall accuracy of ∼20 m (95% confidence level), with a global coverage between 83°N and 83°S (Reuter et al., 2009; Zhao et al., 2011). The most recent and advanced version of GDEM (version 3) has been released since 2019.

4.3.3 Copernicus DEM (GLO-30)

Copernicus DEM emanated from the WorldDEM data which is based on data acquired during the TanDEM-X Mission (Airbus, 2020a). The global accuracy assessment of Copernicus DEM considered about 11 million valid ICESat-GLAS reference points (Airbus, 2020a). The Copernicus GLO-30 dataset has a grid spacing of 1 arc second (30 m) and is available at a standardised extent of 1° x 1°. The DEM is available in two different formats - DGED and DTED. For this study, the DGED format was adopted. More information about the production process of Copernicus DEM can be found in the Copernicus DEM product handbook and validation report (Airbus, 2020a, 2020b).

4.3.4 ALOS World 3D (AW3D) DEM v3.2

The 30 m ALOS World 3D (AW3D30) DEM was released by the Japan Aerospace Exploration Agency (JAXA), and is a derivative of the earlier high-resolution 5 m ALOS DEM (JAXA, 2023). AW3D30 has undergone several improvements since its first release including void filling and reduction in absolute offsets (JAXA, 2023; Okolie & Arungwa, 2022). The most recent version 3.2 was adopted in this research.

4.3.5 FABDEM

The Forest And Buildings removed Copernicus DEM (FABDEM) is a worldwide 30 m elevation map developed by researchers at the University of Bristol and Fathom, an organisation that specialises in water risk intelligence (Hawker et al., 2022). To produce FABDEM, the team developed a correction algorithm using random forest trained on datasets spanning a wide range of urban extents and climatic zones. The result was a globally corrected DEM in which the effect of buildings and trees were minimised. FABDEM is reputedly more reliable than existing global DEMs (Hawker et al., 2022). Due to its recent release, extensive and independent accuracy assessments are required to validate the accuracy of FABDEM. In this research, the accuracy of the globally corrected FABDEM is compared with the enhanced DEMs during performance evaluation.

4.3.6 City of Cape Town (CCT) airborne LiDAR DEM

The City of Cape Town (CCT) airborne LiDAR-derived DEM is a bare-earth digital elevation model of the Cape Town municipal area, acquired from the Information and Knowledge Management Department of the City of Cape Town (City Maps Office). The Department produces annually updated DEMs at 2 m, 5 m and 10 m spatial resolution. All three DEMs are generated independently from LiDAR point clouds, and are produced at a maintenance cycle of 3 years. The point density varies depending on the year. For data acquired prior to 2021, the point density is 2 – 3 points/m^2, but this has been upgraded to 10 points/m^2 post-2021. The height accuracy of the point cloud is 15 cm. The data acquisition for the dataset used in this study was conducted in phases between 2018 and 2021. The DEMs are available in both point cloud and grid/raster formats. The gridded format is derived through processing of the ground-classified LiDAR point clouds using triangulation in LP360, an advanced desktop LiDAR software package. The DEM is spatially referenced to the Hartebeesthoek94 horizontal coordinate system. The height reference used is the South Africa (SA) Land Levelling Datum (LLD).

4.3.7 South Africa national DEM

The 25 m South Africa National DEM was acquired from the Chief Directorate: National Geospatial Information (CD: NGI). The DEM is produced from vectorised contours, or from photogrammetrically sourced data (CD: NGI, 2019). The maintenance of the standard for DEMs is the responsibility of CD: NGI. The directorate provides the following technical specifications to guide the DEM production (CD: NGI, 2019):

1. The absolute height error shall be 1.2m (σ) in flatter areas, and 2m (σ) in steeper areas of South Africa.
2. The horizontal coordinates of the DEM shall be referenced to South Africa's official geodetic datum, Hartebeesthoek4 datum, and the coordinates shall be based upon the Gauss Conform projection.
3. The heights shall be referenced to mean sea level (MSL).

Additional specifications such as the output format, naming convention for DEM tiles, and elevation data capture requirements are also provided in the technical specifications. Malindi (2022) identified several limitations of the SA National DEM to include:

1. Incomplete and inconsistent coverage over South Africa.
2. Missing tiles and missing pixels.
3. The presence of pixels with incorrect height information.

Malindi (2022) also reported that the SA National DEM had a standard deviation of approximately ±4.60 m and ±5.88 m when validated with LiDAR data.

Table 4.1 shows the characteristics of the principal DEMs used in this study while Table 4.2 shows the tile identifiers (IDs) of the DEMs.

Table 4.1: Characteristics of the principal DEMs used in this study

Dataset	Sensor (Source)	Coverage	Horizontal datum	Vertical datum	Height system	Precision	Data acquisition period	Release date of most recent version
NASADEM	C-band radar (NASA)	56°S – 60°N	WGS84	EGM96	Orthometric	Integer	2000	2020
ASTER GDEM	Optical/stereo near infra-red imagery (NASA/JAXA)	83°S – 83°N	WGS84	EGM96	Orthometric	Integer	2000 – 2013	2019
Copernicus GLO-30	X band radar (ESA/Airbus)	Entire Earth	WGS84	EGM2008	Orthometric	Floating point	2010 – 2015	2021
ALOS World 3D	Optical/ Stereo pan imagery (JAXA)	82°S – 82°N	WGS84	EGM96	Orthometric	Integer	2006 – 2011	2022
City of Cape Town (CCT) airborne LiDAR-derived DEM	Information and Knowledge Management Division, City of Cape Town	Cape Town Municipal-ity	Hartebeest-hoek94	SAGEOID10 (Land Levelling Datum, LLD)	Orthometric	Floating point	2018 – 2021	-

Sources: Author's compilation; Takaku et al. 2015; Guth and Geoffroy 2021; Hawker et al. 2019; Purinton and Bookhagen 2021; Airbus, 2020a, 2020b

Table 4.2: Tile IDs of the DEMs

DEM	Tile ID
NASADEM	s34e018 s35e018
ASTER	ASTGTMV003_S34E018 ASTGTMV003_S35E018
Copernicus	Copernicus_DSM_10_S34_00_E018 Copernicus_DSM_10_S35_00_E018
AW3D	ALPSMLC30_S034E018 ALPSMLC30_S035E018
FABDEM	S34E018_FABDEM_V1-2 S35E018_FABDEM_V1-2
South Africa National DEM	LO19_050M_3318CB LO19_050M_3318CD LO19_050M_3318DA LO19_050M_3318DB LO19_050M_3318DC LO19_050M_3318DD LO19_050M_3418AB LO19_050M_3418AD LO19_050M_3418BA LO19_050M_3418BB LO19_050M_3418BD LO19_050M_3419AA LO19_050M_3419AC

4.3.8 Global tree cover and bare ground

The global tree cover (treecover2010) and bare ground cover (bareground2010) provide estimates of the percent maximum tree canopy cover and the percent bare ground cover respectively (GLAD, 2023a; GLAD, 2023b). To produce the datasets, regression models estimating the per pixel cover were applied to Landsat imagery (Hansen et al., 2013). In the Landsat analysis for treecover2010 which was carried out in Earth Engine, trees were defined as all vegetation taller than 5 m (Hansen et al., 2013). The methodology is extensively discussed in Hansen et al. (2013).

4.3.9 Global urban footprint

The Global Urban Footprint campaign was introduced by the German Aerospace Centre (DLR) (Esch et al., 2012, 2018). GUF is a database of human settlements derived from the TanDEM-X (TDM) SAR data. Its production is based on the Urban Footprint Processor (UFP) processing environment (Esch et al., 2013; Esch et al., 2017). The methodology for deriving the GUF is presented in Esch et al. (2010, 2013). Due to a variety of challenges, the GUF settlement

classifications are subject to varying accuracies across global landscapes (Mück et al., 2017). The 0.4 arc-second (~12 m) 2012 GUF dataset was adopted in this research.

4.3.10 Terrain parameters

Terrain heterogeneity and variability is an important variable for understanding the spatial variation of DEM error. There are a vast array of terrain parameters that are useful for topographic and DEM analysis, and that are important for quantifying terrain variability within a certain area. In this study, the following terrain parameters are considered: slope, aspect, surface roughness, topographic position index (TPI), terrain ruggedness index (TRI), terrain surface texture (TST), and vector ruggedness measure (VRM).

The theoretical underpinnings of these parameters are already discussed in Chapter 2. The terrain parameters were generated within the QGIS 3.28.2 and SAGA GIS 7.8.2 software environments. QGIS supports the creation, visualisation, analysis and publishing of geospatial information (QGIS, 2023). The System for Automated Geoscientific Analyses (SAGA) has a comprehensive set of tools for implementation of spatial algorithms. Both QGIS and SAGA GIS are free and open-source.

4.4 Mosaicking

Mosaicking involves the creation of large area DEMs from multiple input/source DEMs of the same spatial reference, or the compositing of DEMs for the same area by averaging height values within areas of overlap (Knöpfle et al., 1998; Okolie & Smit, 2022). The tiles for each DEM product shown in Table 4.2 were mosaicked within the ArcGIS environment (Figure 4.4). In the mosaic procedure, one of the DEM tiles is selected as a target raster (i.e., the DEM tile to which the other DEM tiles will be added) and the ‘first’ mosaic operator was adopted (i.e., The output pixel height value of the overlapping areas will be the value from the first DEM tile mosaicked into that location).

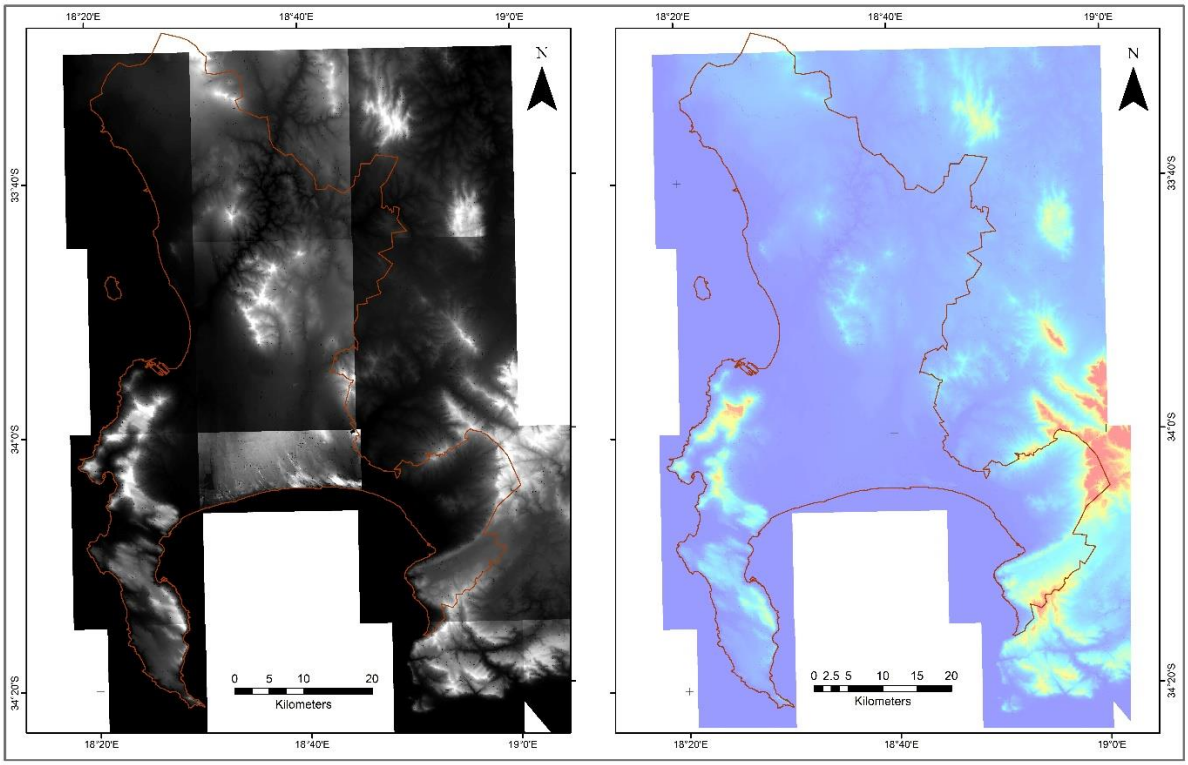


Figure 4.4: South Africa national DEM tiles before (left) and after being mosaicked (right)

4.5 Datum Harmonisation

4.5.1 Horizontal datum transformation

To achieve uniformity in the spatial reference systems, the DEMs were projected from the 3D Geographic to the Universal Transverse Mercator (UTM) coordinate system (UTM Zone 34 – southern hemisphere). It is generally understood that map projections incur some level of distortion during the process of transforming 3D spherical coordinates onto a plane (Laurini & Thompson, 1992). However, the UTM projection has been adopted as an international standard. It is widely used in topographic maps, and is recommended for areas on the earth lying between latitudes 84°N to 80°S (Emery & Camps, 2017; Jain & Singh, 2003). UTM provides a constant distance relationship and minimises projection distortions. More so, the projection is a necessary step to convert the DEM horizontal coordinates to planimetric units before deriving terrain parameters such as slope. The global DEMs were transformed from the geographic to the UTM WGS84, and the Cape Town LiDAR DEM was transformed from Hartebeesthoek94 to UTM WGS84. Similarly, the South Africa National DEM was transformed to the UTM system.

4.5.2 Vertical datum transformation

Compared to EGM96, the development of EGM2008 followed a more rigorous approach and better algorithms (Roman et al., 2010). For example, EGM2008 leads to improved estimation

of topographical heights (see Pavlis et al., 2012; Üstün et al., 2016). The root mean square difference between EGM96 and EGM2008 is about 0.5 m, and this could reach up to 10 m in areas such as ocean trenches (USNA, 2021). In a numerical investigation in Turkey, Üstün et al. (2016) showed that a conversion from EGM96 to EGM2008 could improve the height accuracy by up to 2.5 m. In external validations of EGM2008 over parts of Africa, e.g., South Africa (Merry, 2009) and Algeria (Benahmed Daho, 2009) it provided a better agreement than EGM96 when compared to global positioning system (GPS)/levelling data. The review of previous studies favours the adoption of EGM2008 as a common vertical datum for the DEMs in this study. Moreover, the use of EGM2008 in topographic mapping projects has been encouraged (e.g., Rao et al., 2012). Figure 4.5 illustrates the derivation of elevations using geoidal heights estimated from EGM96 and EGM2008.

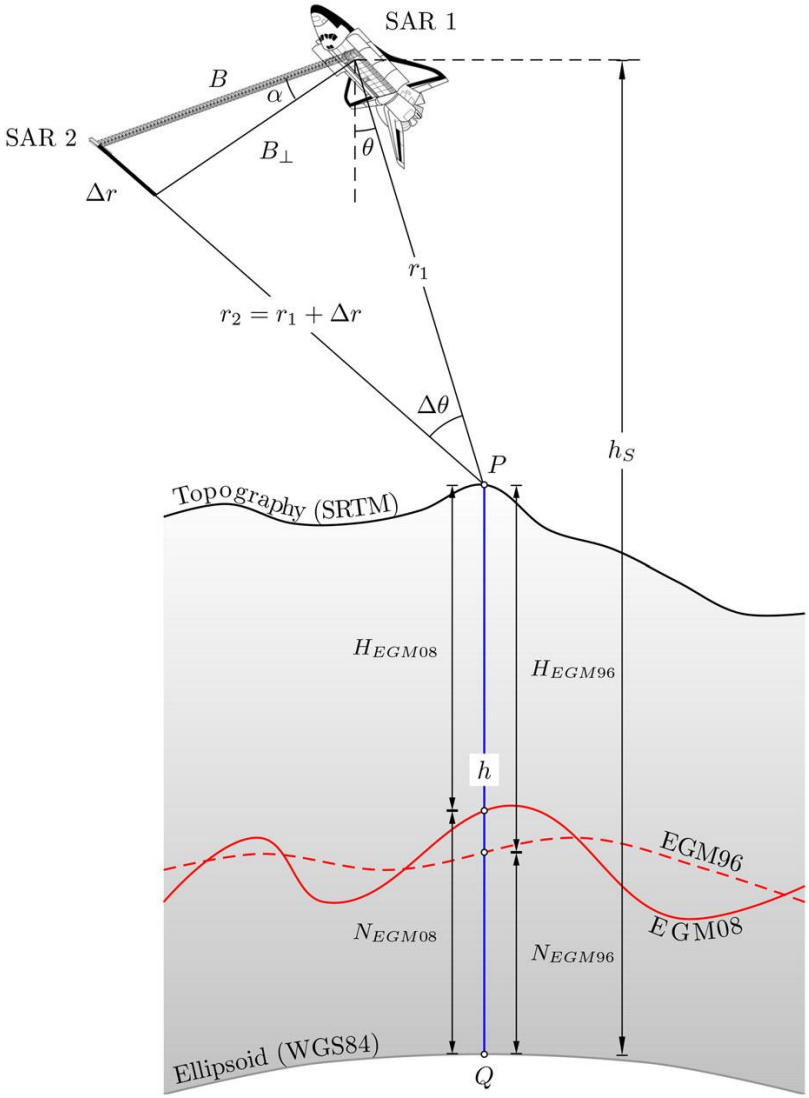


Figure 4.5: Derivation of elevations above sea level using geoidal heights estimated from EGM96 and EGM2008 illustrated here with geometrical aspects of SRTM, from which NASADEM originated (Source: Üstün et al., 2016, p. 27)

The EGM96 and EGM2008 geoid models were downloaded from the website of the International Centre for Global Earth Models (ICGEM). The following parameters were defined for retrieving the geoid models through ICGEM online (http://icgem.gfz-potsdam.de/calcgrid; Figure 4.6):

1. Calculation Service: Regular grids
2. Model selection: (i) EGM96 (ii) EGM2008
3. Functional selection: geoid
4. Grid selection: 18.26 – 19.03E; 33.45 – 34.38S
5. Grid step: 0.0005 deg
6. Reference System: WGS84

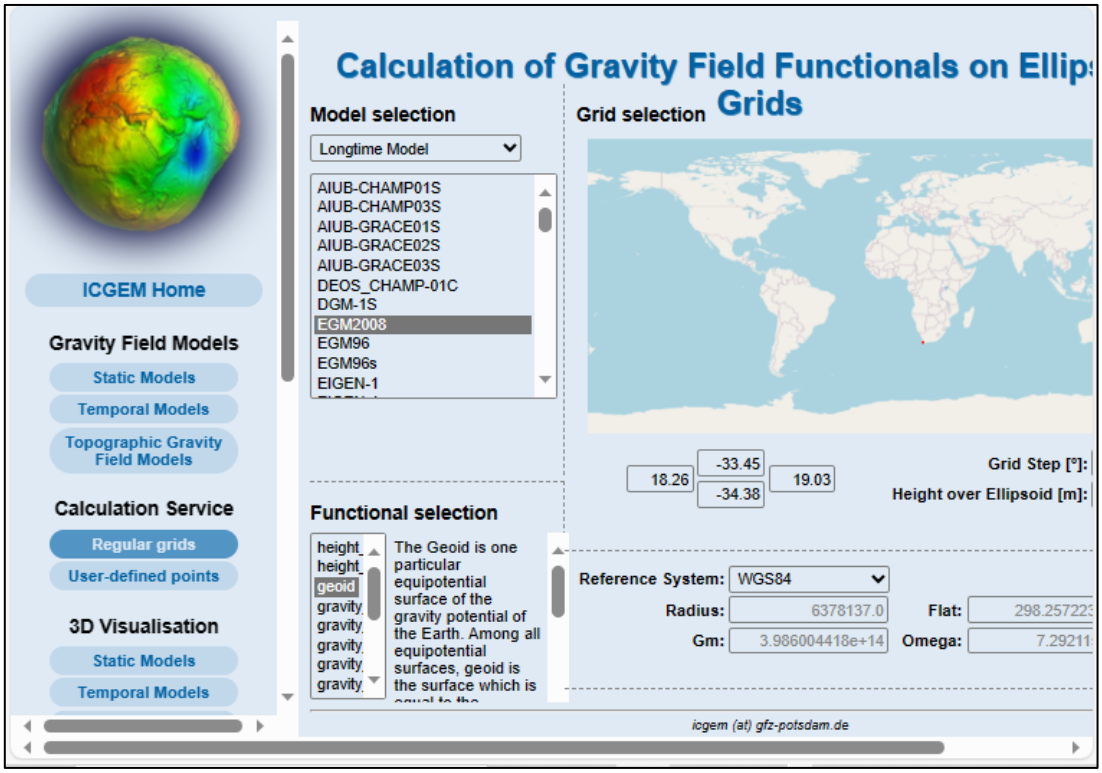


Figure 4.6: The EGM96 and EGM2008 geoid models were retrieved from ICGEM

The geoid heights were retrieved in plain text format at a grid spacing of 0.0005°. The grids were converted to raster surfaces through inverse distance weighted (IDW) interpolation (Figure 4.7) in ArcGIS 10.8, and the differences in the geoid heights were calculated.

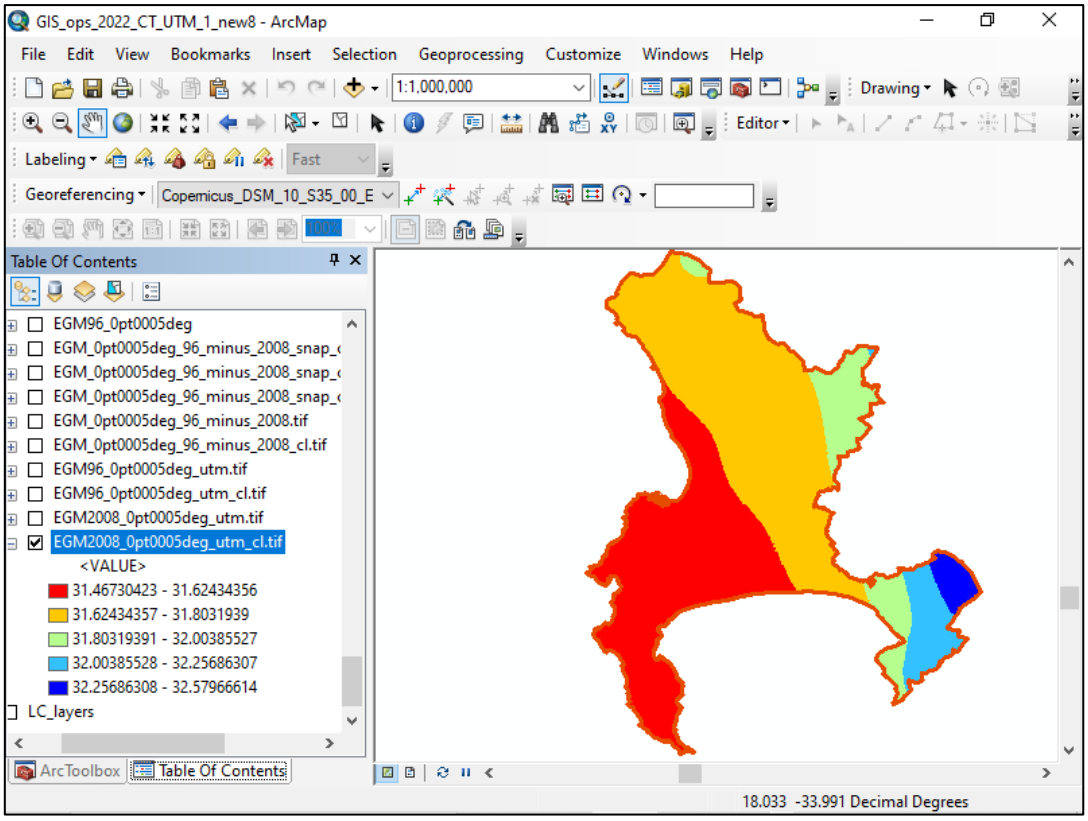


Figure 4.7: The EGM grid points were rasterised through IDW interpolation

Thereafter, the vertical datum conversion from EGM96 to EGM2008 was carried out within the ArcGIS raster calculator, using the equation by Üstün et al. (2016).

$$H_{EGM08} = H_{EGM96} + N_{EGM96} - N_{EGM08} = H_{EGM96} + \delta N \quad (4.1)$$

Where,

H_{EGM08} – elevation value from global DEM (EGM2008)

H_{EGM96} – elevation value from global DEM (EGM96)

N_{EGM08} - geoid height (EGM2008)

N_{EGM96} - geoid height (EGM96)

δN - geoid height difference between EGM96 and EGM2008

The South African hybrid geoid model (SAGEOID10) was developed by Chandler & Merry (2010) for ellipsoidal to orthometric height transformation on the South African Land Levelling Datum (LLD) (Mphuthi, 2021). Thus, the SAGEOID10 geoid model was acquired from South Africa's national mapping agency, CD: NGI. This data is stored in ASCII format, on a 2.5' grid that is compatible with EGM2008 data. Both data points (SAGEOID10 and EGM2008) were interpolated to rasters (gridded format) with a cell size matching the LiDAR DEM. Finally, the vertical datum conversion to EGM2008 was implemented as follows:

$$H_{EGM08} = H_{SAGEOID10} + N_{SAGEOID10} - N_{EGM08} = H_{SAGEOID10} + \delta N \quad (4.2)$$

Where,

H_{EGM08} - elevation value from LiDAR DEM (EGM2008)

$H_{SAGEOID10}$ - elevation value from LiDAR DEM (SAGEOID10, South African LLD)

N_{EGM08} - geoid height (EGM2008)

$N_{SAGEOID10}$ - geoid height (SAGEOID10, South African LLD)

δN - geoid height difference between SAGEOID10 (South African LLD) and EGM2008

Figure 4.8 shows a comparison of the EGM96 and EGM2008 geoid models, while Figure 4.9 shows a comparison of the SAGEOID10 and EGM2008 geoid models. Following the same approach, the South Africa National DEM was transformed to EGM2008.

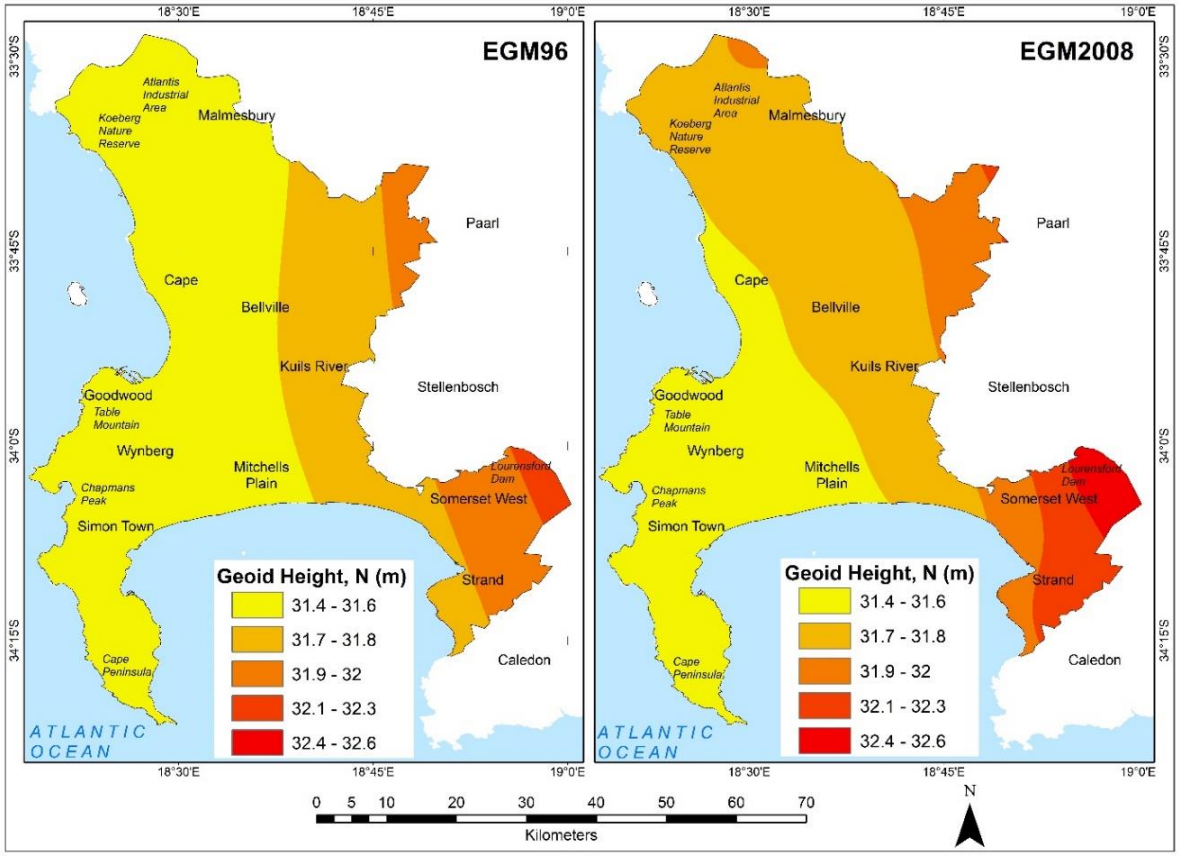


Figure 4.8: EGM96 and EGM2008 geoid models for Cape Town

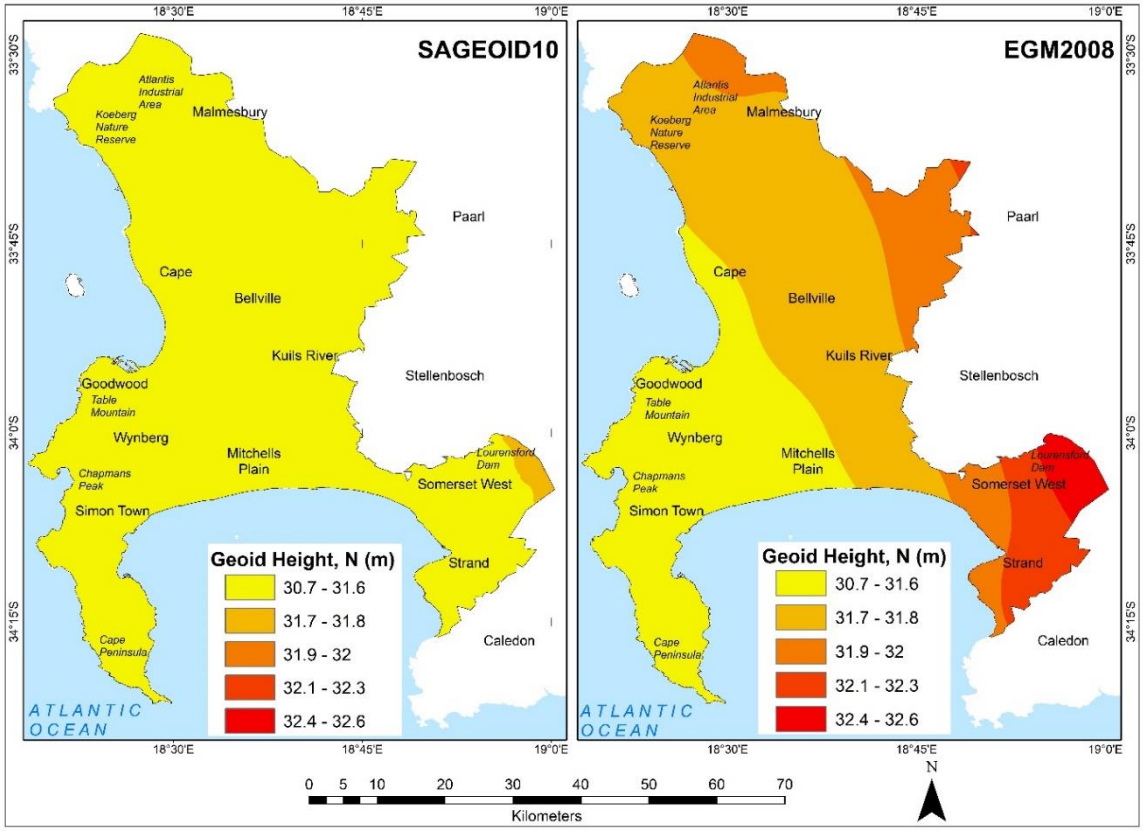


Figure 4.9: SAGEOID10 and EGM2008 geoid models for Cape Town

4.6 Quality Assessment and DEM Selection

The vertical accuracy of DEMs is influenced by several factors such as the nature of the terrain. An adequate accuracy assessment of DEMs should consider several factors including the spatial distribution of ground control points, DEMs of difference and terrain derivatives (Höhle & Höhle 2009; Carrera-Hernández, 2021). In this section, qualitative and quantitative analytical approaches are combined for assessing the vertical accuracy of the four global DEMs using the high-resolution 2 m airborne LiDAR as a reference. The accuracy of checkpoints used for validating a DEM should be (Höhle & Höhle, 2009; Maune, 2007):

1. At least three times more accurate than the DEM elevations being evaluated,
2. Distributed randomly,
3. Sufficiently large in number.

The method adopted for quality assessment also depends on the available reference data (see Iheaturu et al., 2020).

4.6.1 Methodological framework for quality assessment and DEM selection

Figure 4.10 presents a workflow diagram of the methodology for quality assessment.

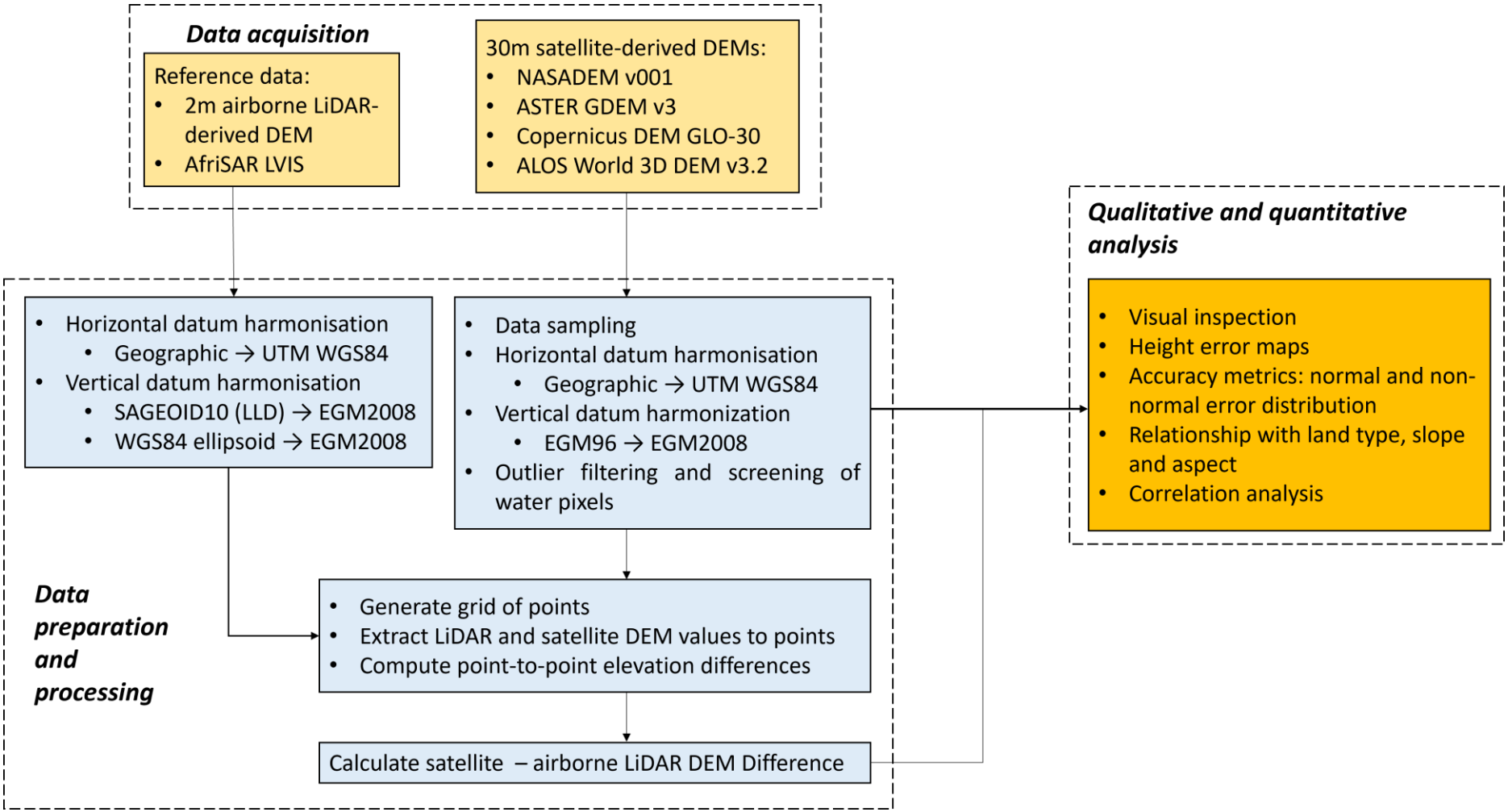


Figure 4.10: Workflow diagram of the DEM quality assessment methodology

4.6.2 Selected sites and data distribution

The whole area of Cape Town was considered for the DEM quality assessment. In addition, the five landscapes described in section 4.2 were considered for a more detailed assessment. The specific sites chosen for this stage are shown in Figure 4.2. These sites are topographically and geomorphologically diverse. More importantly, the sites encompass different terrain conditions coexisting with various land cover. Table 4.3 presents the general characteristics of these specific sites.

Table 4.3: General site characteristics of landscapes for the DEM quality assessment

Landscape	Description
Whole area	Whole area of Cape Town with mixture of landscapes.
Urban/ industrial	Densely built-up residential, commercial and industrial area in Cape Town Municipality.
Agricultural	Agricultural fields and few farm settlements. Meeting point of Diep River and Mosselbank River.
Grassland/ shrubland/ dense bush	Light vegetation, shrubs, open grassland and dense bushes, southeast of Grotto Bay.
Mountain	Flat-topped Table Mountain National Park, with the adjoining Camps Bay and Happy Valley.
Peninsula	Low-lying and gently sloping section of the Cape Peninsula, in the general vicinity of Olifantsbos Beach, Miller's Point and Partridge Point.

4.6.3 Data preparation

In this assessment, a grid of 2,760,581 data points was created. Subsequently, elevations from the LiDAR DEM and the global DEMs that coincided with the data points were extracted using the point-to-raster tool in ArcGIS (ESRI, 2021a). Polidori & Hage (2020) have stated that "the goal of the external validation is to assess the quality of a DEM using external reference data, which can be a sparse cloud of points." Moreover, point data are the most frequently used in DEM validation studies (Mesa-Mingorance & Ariza-López, 2020).

4.6.4 Outlier filtering and screening of water pixels

Outliers in the elevation data from the global DEMs were filtered in three stages. Firstly, negative elevation values were deleted. After this, the Three-sigma rule (Wessel et al. 2018) for outlier filtering was applied whereby values outside the mean elevation difference (Δh_{mean}) plus/minus three times the standard deviation ($\Delta h > \Delta h_{mean} + 3 * SD$, $\Delta h < \Delta h_{mean} - 3 * SD$) were eliminated. In the third stage, elevation points falling on water bodies were also eliminated. Specular reflection and temporal decorrelation in InSAR acquisitions and

processing could cause smooth water surfaces to appear as rough surfaces (Ettritch et al., 2018; Huber et al., 2021). Thus, elevation points that intersected with natural and artificial water bodies (identified from Google Earth imagery and a Cape Town waterbodies dataset) were removed. This resulted in a final set of 2,570,126 points for validation of the DEMs.

4.6.5 Accuracy assessment

Quantitative and qualitative (visual) analysis are standard checks used as quality assessment criteria for DEMs (Gesch et al., 2014; Cenci et al., 2021). Powerful visual analysis can reveal the existence of problematic areas (Mesa-Mingorance & Ariza-López, 2020). To determine the absolute vertical accuracy, the difference between the global DEMs and the reference DEM (LiDAR) was calculated. To derive the elevation errors (ΔH), the LiDAR elevations were subtracted from the corresponding elevations of the global DEMs at specific points. That is, the elevation values from the airborne LiDAR and satellite DEMs were extracted wherever they intersected at specified grid points. This analysis of the elevation difference (ΔH) enables the detection of possible random errors and systematic bias (Cenci et al., 2021).

The elevation differences (ΔH) between the DEMs and reference LIDAR were calculated:

$$\Delta H = H_{GlobalDEM} - H_{RefDEM} \quad (4.3)$$

Where,

H_{RefDEM} = elevation from LiDAR, and $H_{GlobalDEM}$ = individual elevations from NASADEM, ASTER, Copernicus and AW3D

Correlation analysis enabled the analysis of the inter-relationships between the elevation errors. Using the statistical populations of the ΔH and/or |ΔH|, several accuracy metrics were computed as follows: mean error (ME), standard deviation (SD), root mean square error (RMSE), mean absolute error (MAE), median absolute deviation (MAD), normalised median absolute deviation (NMAD) and linear error at 95% (LE95) confidence level. The RMSE is generally the preferred metric in DEM accuracy studies. However, the RMSE also has limitations as an accuracy measure. For example, it does not distinguish between random errors, systematic errors or blunders; and the errors in a DEM do not always follow a normal distribution (Wise 2000; Höhle & Höhle, 2009; Carrera-Hernández, 2021). There have been several debates and empirical arguments over which of the metrics (MAE or RMSE) should be preferred (e.g., (Brassington, 2017; Chai & Draxler, 2014; Christie & Neill, 2022; Schneider

& Xhafa, 2022). A standard practice has been to calculate both metrics for a balanced perspective. Nonetheless, the RMSE is an excellent error metric for evaluating the accuracy of numerical predictions (Christie & Neill, 2022). The robust MAD and NMAD overcome these limitations (Willmott & Matsuura, 2005; Höhle & Höhle, 2009; Carrera-Hernández, 2021). The equations are given as follows:

$$\mathrm{ME} = \frac{\sum_{i=1}^{i=n}(\Delta H)}{n} \tag{4.4}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{i=n}(|\Delta H|)}{n} \tag{4.5}$$

$$\mathrm{SD} = \sqrt{\sum_{i=1}^{n} \frac{(\Delta H_i - \underline{\Delta H})^2}{n-1}} \tag{4.6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\Delta H_i{}^2\right)} \tag{4.7}$$

Where,

$\underline{\Delta H}$ = mean of the height differences

The MAD and NMAD were calculated as follows (Höhle & Höhle, 2009; Wessel et al., 2018):

$$\mathrm{MAD} = \mathrm{MD}\left(\left|(\Delta H_i - \mathrm{MD}_{\Delta H_i}\right|\right) \tag{4.8}$$

$$\mathrm{NMAD} = 1.4826 \times \mathrm{MAD} \tag{4.9}$$

At confidence level of 95%, the LE is calculated as given below:

$$\mathrm{LE95} = \mathrm{SD} \times 1.96 \tag{4.10}$$

Where,

n is the number of elevation point samples

MD is the median of a set of data

$\mathrm{MD}_{\Delta H_i}$ is the median of the elevation differences or the 50% quartile.

The descriptive statistics were summarised using SPSS and Microsoft Excel, while statistical visualisations were generated with R software.

4.6.6 Height error maps

Height error maps (HEMs), were generated by subtracting the LiDAR DEM from the global DEMs, as a visual estimation of the random height error in different landscapes (Gdulová et al., 2020). Because the LiDAR elevation pixels that intersected or coincided with the 30 m gridded points were extracted, there was no need to resample the LiDAR. According to Hawker et al. (2019), HEMs are useful for visualising the spatial distribution of elevation errors and enable an indication of the error sources.

4.6.7 Spatial pattern of DEM error

The spatial pattern of DEM error was evaluated using the Global Moran's I tool in ArcGIS (Figure 4.11). Essentially, the Global Moran's I tool evaluates characteristics of the spatial pattern (i.e., random, dispersed or clustered) (ESRI, 2023).

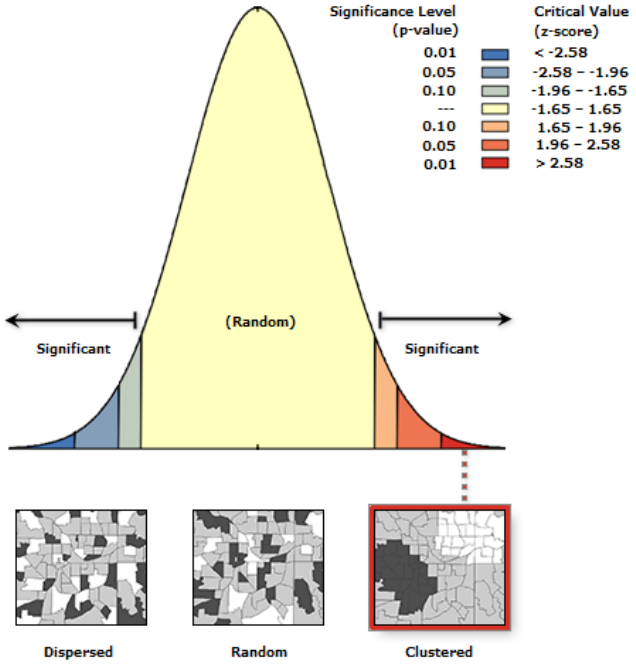


Figure 4.11: Visualised output of spatial pattern analysis (ESRI, 2021, para. 3)

4.6.8 Geo-morphometric analysis

The next analysis involved an assessment of the relationship between vertical error and two primary terrain parameters, slope and aspect. The vertical error of DEMs and their derived parameters are of great importance since the vertical errors can be propagated to the terrain parameters (e.g., Nwilo et al. 2022). This error propagation could also lead to erroneous watershed delineation or drainage networks (Wechsler, 2007; Hengl et al., 2010; Polidori & El Hage, 2020). The guidance for slope classification was available from various sources (e.g., Canadian Soil Information Service, 2013; Chesworth et al., 2008). For this analysis, the slope

and aspect maps derived from the LiDAR DEM were grouped into six slope classes and eight cardinal directions respectively (Figure 4.12). Thereafter, the elevation errors were grouped based on these classes.

Going further, the potential of the DEMs for geo-morphometric modelling and analysis based on the achieved vertical accuracies was evaluated. Following the approach of Acharya et al. (2018), Barber & Shortridge (2013) and Podhorányi et al. (2013), the LiDAR DEM was downsampled from 2 m to a lower resolution of 30 m. This provides a good approximation of the original ground surface while enabling a fair comparison with the satellite DEMs. Refining the satellite DEMs through upsampling might be inappropriate due to the inability of obtaining new information (Podhorányi et al., 2013). Moreover, for regional-scale analysis, use of high-resolution data may impose constraints in software capability, processing time and available memory (Grohmann, 2015). Next, a sub-basin was selected for analysis (Figure 4.13). After some pre-processing to fill sinks in the DEMs, several geo-morphometric parameters were generated for comparison using the morphometric toolbox (Beg, 2015, 2020). The list of computed parameters is shown in Table 4.4.

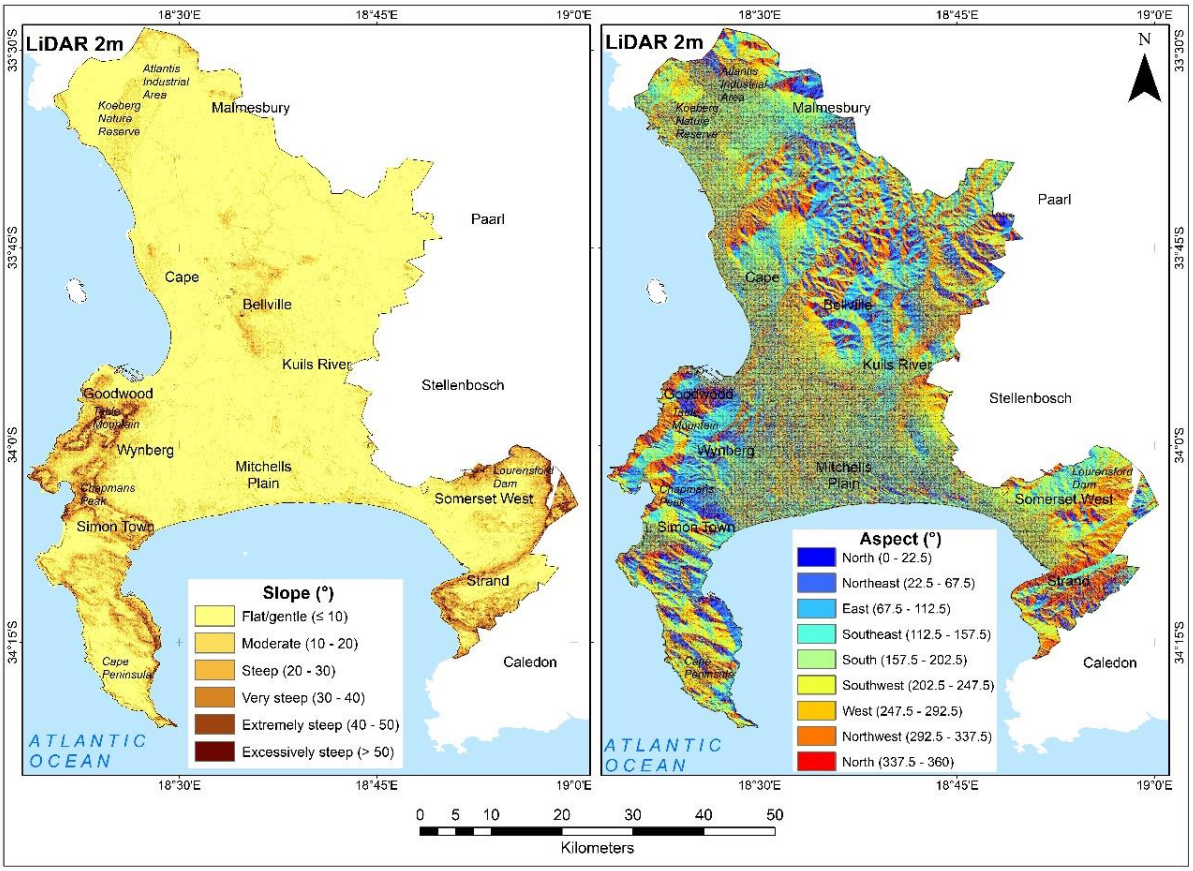


Figure 4.12: Slope (left) and aspect (right) classification using airborne LiDAR

Table 4.4: Description of the geomorphometric parameters

S/N	Geo-morphometric Parameter	Description
1	Stream order	The level of branching in a river system (Bogale, 2021)
2	Total Basin Area	The projected area enclosed by the basin boundary (Horton, 1945)
3	Basin Length	This is the distance from outlet to the farthest point on the basin boundary (Nwilo et al., 2021; Schumm, 1956)
4	Main Channel Length	The length of longest water course from the outlet to upstream
5	Mean Basin Width	The total basin area divided by the basin length (Horton, 1932)
6	Drainage Texture	The total number of stream segments of all orders per perimeter of that area (Dikpal et al., 2017; Horton, 1945)
7	Drainage Density	The intricate arrangement of stream networks. A higher drainage density value implies higher runoff (Horton, 1945; Tesema, 2021; Prasad et al., 2008)
8	Stream Frequency	The total number of stream segments of all orders per unit area (Dikpal et al., 2017; Horton, 1932, 1945)
9	Drainage Intensity	The stream frequency divided by the drainage density (Tesema, 2021)
10	Length of Overland Flow	The length of water over the ground surface before getting concentrated into certain definite stream channels. (Horton, 1945; Sukristiyanti et al., 2018)
11	Height of Basin Outlet (m)	Selected point elevation from DEM
12	Maximum Height of Basin (m)	Selected point elevation from DEM
13	Total Basin Relief (H) (m)	Maximum relief – Minimum Relief (Horton, 1945)

4.7 Co-registration and Resampling

Using a combination of the projection tools and the snap raster environment setting in ArcGIS, the four global DEMs were projected from geographic to UTM coordinate system (WGS84) and harmonised to same cell size and alignment of a 30 m reference grid. In the snap raster operation, the lower left corners of the four DEMs were moved to the nearest reference grid cell corner while the upper right corner of the DEMs were moved to the nearest output reference grid cell corner. As such, the cells in the four global DEMs were aligned such that their cell centres coincided.

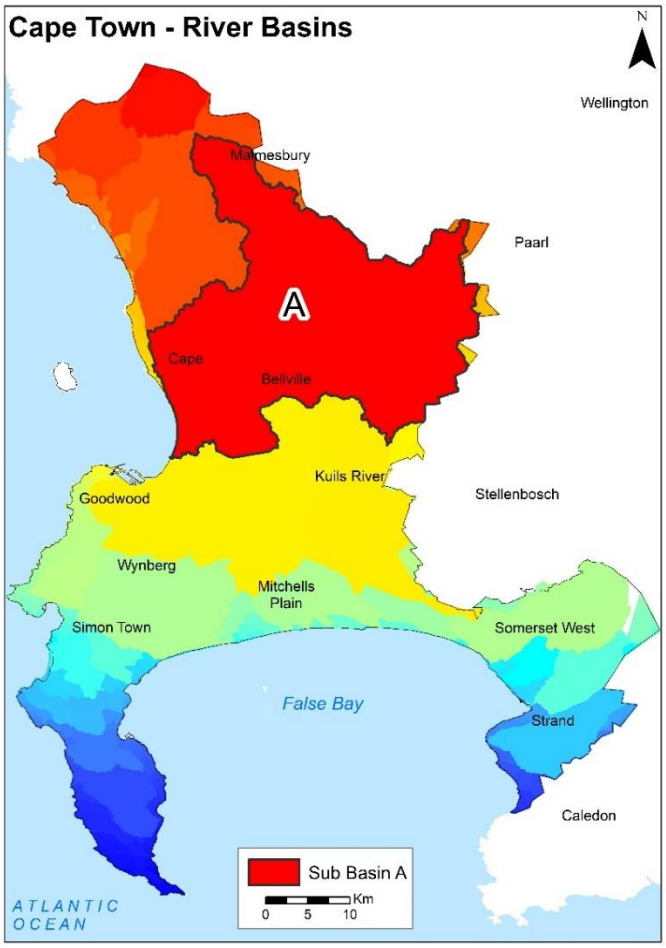


Figure 4.13: River basins showing the selected sub-basin for drainage network analysis

4.8 Tree-based Ensembles for DEM Enhancement

In this section, ten tree-based ensemble machine learning algorithms are evaluated for DEM enhancement. Decision trees have a high tolerance for multicollinearity (Climent et al., 2019; Han et al., 2019; Pham & Ho, 2021). In some cases, attempts to remove multicollinearity between features in decision trees could result in poor performance (Kotsiantis, 2013). If there are highly correlated features or variables, decision trees are inherently able to choose only one of the features when deciding upon a split (Climent et al., 2019). Therefore, the present analysis retains all the independent variables because it is not necessary to remove them.

4.8.1 Selected tree-based ensemble algorithms

Ten tree-based ensembles (Table 4.5; Figure 4.14) were evaluated for the enhancement of Copernicus DEM in agricultural lands of Cape Town. The selected algorithms are already discussed in sections 2.8 and 3.4. Several studies have shown a relationship between topography and soil, biomass, organic matter, and crop yields (Ma et al., 2020; Nie et al., 2019; Qi-yong et al., 2014). Thus, improvement of DEM accuracy in agricultural lands is an important concern, and serves as a critical evaluation of the potential of tree-based ensembles.

Table 4.5: Classification of ten tree-based ensemble algorithms considered in this assessment

Bagging: • Random forest • Bagging meta-estimator	**Boosting:** • Adaptive boosting (AdaBoost) **Gradient boosting:** • Gradient boosting machine (GBM) • Extreme gradient boosting (XGBoost) • Light gradient boosting machine (LightGBM) • Histogram-based gradient boosting • Categorial boosting (CatBoost) **Natural gradients:** • Natural gradient boosting (NGBoost)	**Improvisation to gradient boosting:** Regularised greedy forest (RGF)

Figure 4.14: Importing the model regressors

4.8.2 Methodological framework for the evaluation of tree-based ensembles

Figure 4.15 shows the workflow diagram of the assessment methodology adopted. The following DEM (elevation) error conditioning variables were selected: elevation, slope, aspect, surface roughness, topographic position index, terrain ruggedness index, terrain surface texture and vector ruggedness measure.

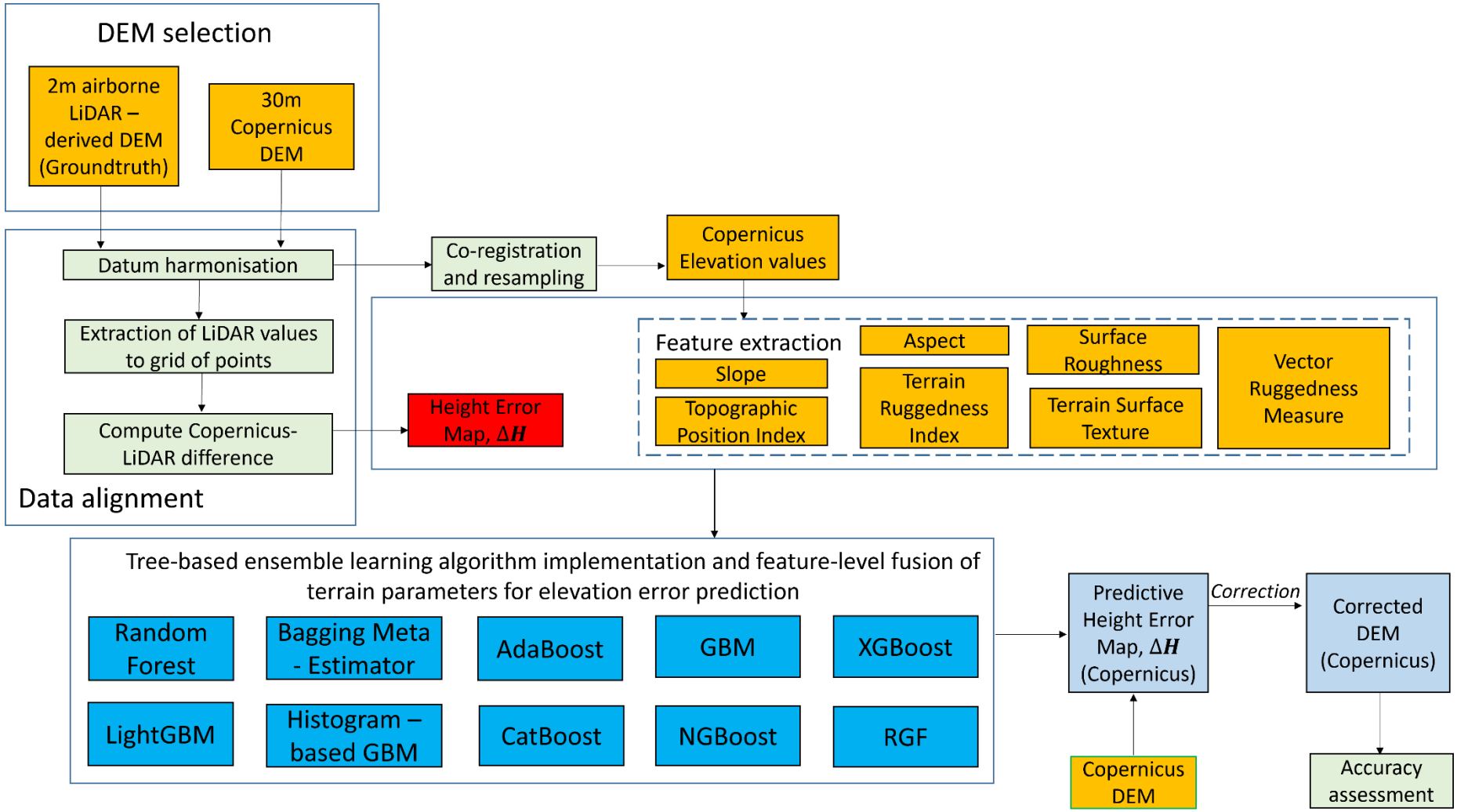


Figure 4.15: Methodological framework for the evaluation of tree-based ensembles

4.8.3 Selected sites and data distribution

The sites (shown in Figure 4.16) are selected from agricultural lands with few settlements located along the floodplain of the Diep River. The Diep River, with its sources in the Riebeek-Kasteel Mountains, flows in a south-west direction towards Table Bay where it widens out forming a large wetland before entering the sea (Drakenstein Municipality, 2009). The Diep River is a sub-catchment within the Berg Catchment area, and is surrounded by informal settlements, industries, factories and farms (DWS, 2020; Gqomfa et al., 2022). The river catchment is low-lying and flat, and falls into the western lowland area of the Western Cape (DWS, 2002). Table 4.6 shows the distribution of sample points at the selected sites. After model training and testing, the models were externally evaluated at separate sites referred to as the 1st, 2nd and 3rd model implementation sites respectively (Figure 4.16, Figure 4.17). The model implementation provides an opportunity for external validation of the trained models. Figure 4.18 shows some attribute characteristics of the training/test dataset used for evaluation of tree-based ensembles.

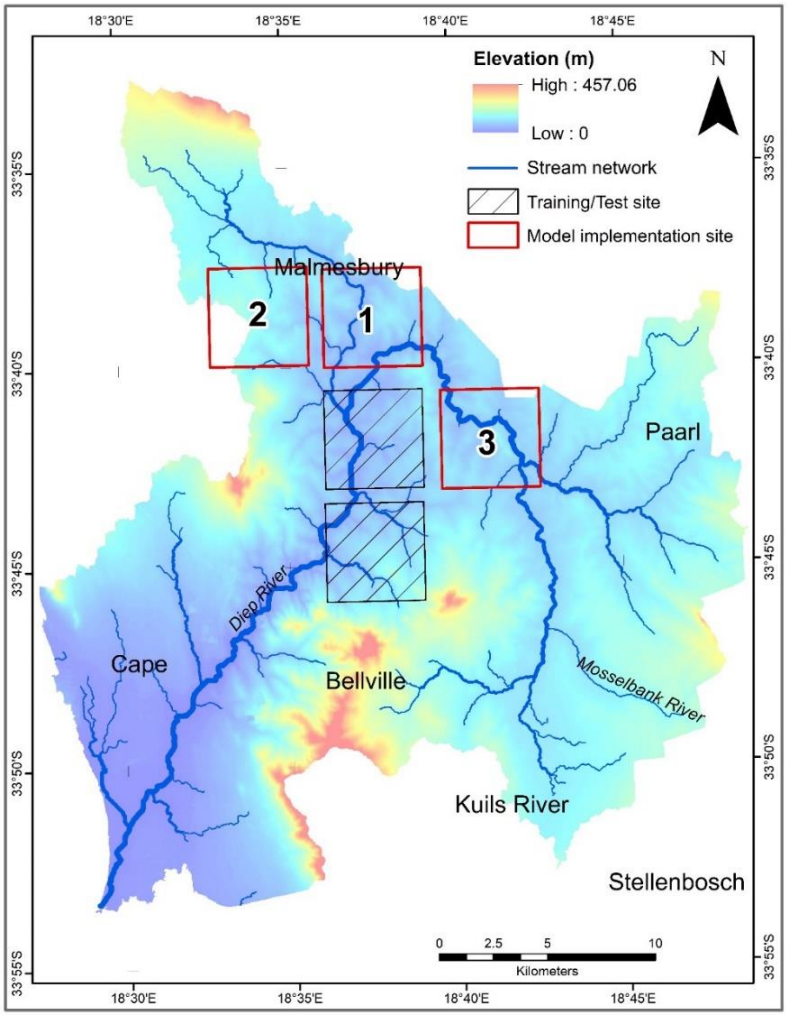


Figure 4.16: Selected sites and data distribution for the evaluation of tree-based ensembles

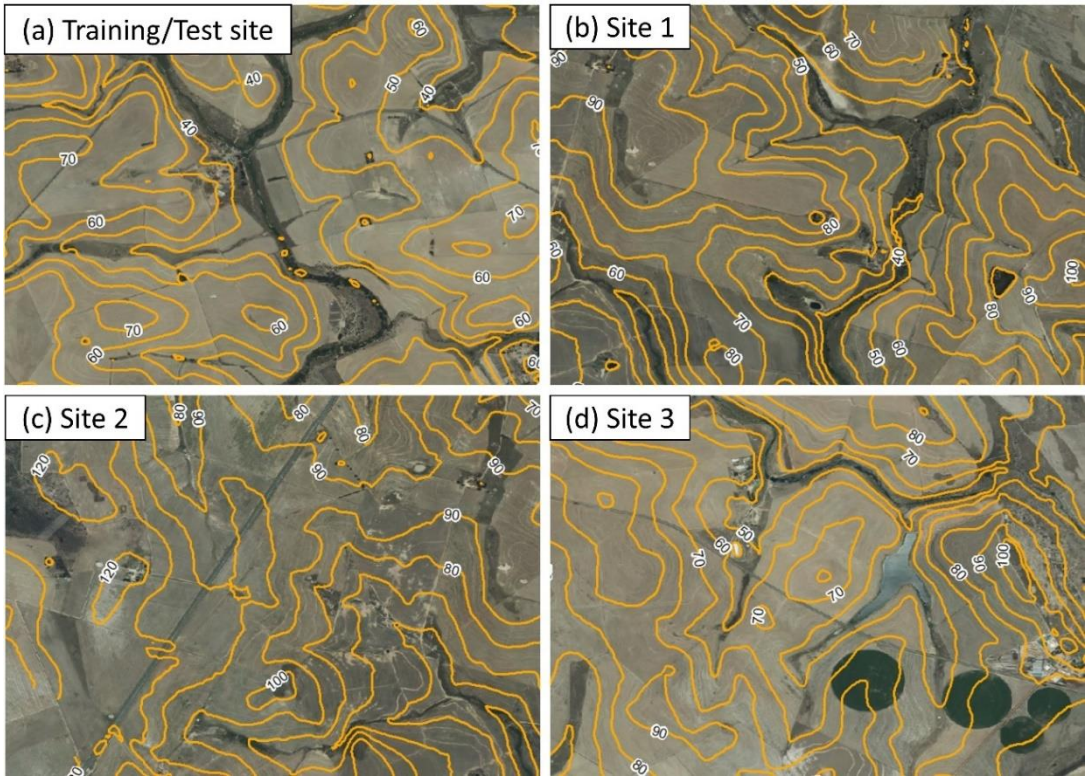


Figure 4.17: View of agricultural lands selected for evaluation of tree-based ensembles – training/test site (top left), 1st model implementation site (top right), 2nd model implementation site (bottom left), 3rd model implementation site (bottom right) (Aerial imagery, January 2023; Source: City of Cape Town). Contour interval: 10 m

Table 4.6: The distribution of sample points at the selected sites

Site ID	No of points	Area (km^2)
Training/testing	46177	41.4
1st implementation site	21539	20.7
2nd implementation site	23109	20.7
3rd implementation site	22366	20.7

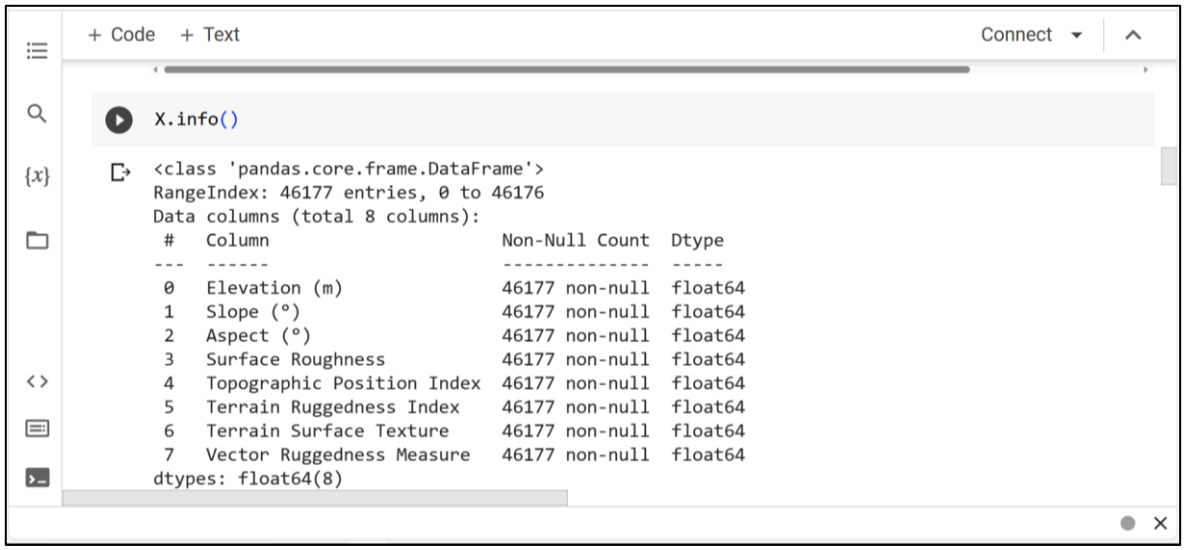


Figure 4.18: Attribute characteristics of the training/test dataset used for evaluation of tree-based ensembles

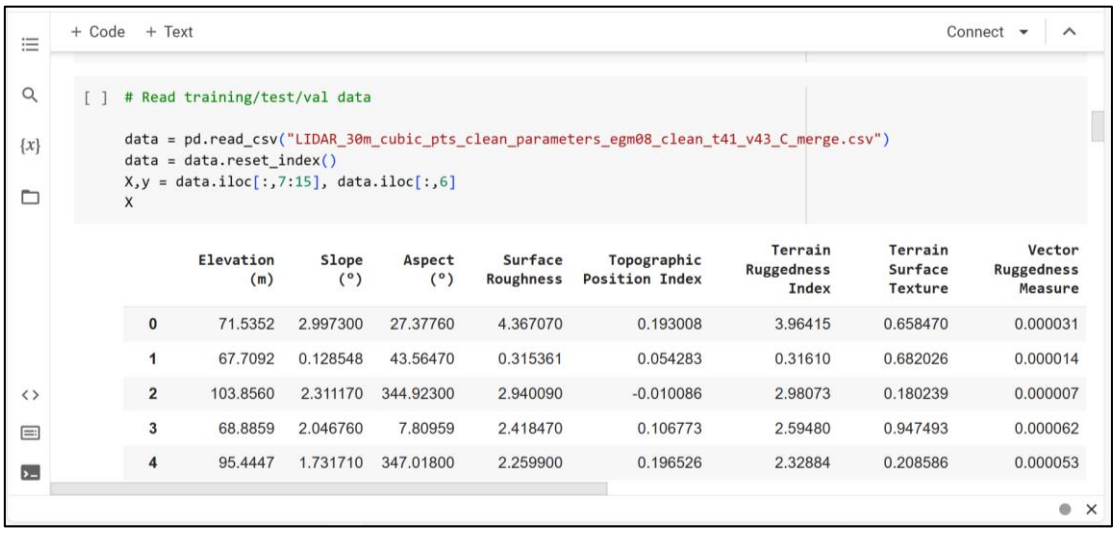
4.8.4 Model implementation and DEM correction

The ten models were implemented using their default hyperparameter values. A hyperparameter is "a parameter whose value is given by the user and used to control the learning process" (Mariani & Sipper, 2022). Their values "control the learning process and determine the values of model parameters that a learning algorithm ends up learning" (Nyuytiymbiy, 2020). The training data includes the elevation and derived terrain parameters (slope, aspect, surface roughness, topographic position index, terrain ruggedness index, terrain surface texture, and vector ruggedness measure) as input parameters, and the elevation error (ΔH) as the target variable. The elevation error was calculated following the procedure in section 4.6.5. The training datasets were converted from raster to csv format. A split ratio of 80% for training and 20% for testing was used. The models were implemented in the Google Collaboratory (Colab) cloud computing environment using Python scripting, the scikit-learn machine learning library and other open-source libraries/packages. Colab enables the writing and execution of Python code through web browsers, and is well suited for machine learning (Google, 2023). The default CPU for the Colab processing environment has the following specifications - Intel Xeon CPU @ 2.20GHz, 13 GB RAM, Tesla K80 accelerator, and 12 GB GDDR5 VRAM (Das, 2022).

A list of Python packages used (e.g., Pandas; Figure 4.19) and their descriptions are presented in Table 4.7. The data was passed into the model regressors for training (e.g., Figure 4.20). Afterwards, the models were implemented for predicting the elevation errors at three implementation sites (shown in Figure 4.17). The predicted errors were subsequently applied for correcting the DEMs (i.e.., $\mathrm{DEM_{Corrected}} = \mathrm{DEM_{Original}} - \Delta\mathrm{H_{Predicted}}$) at the implementation sites, and the corrected data were converted from csv to raster format.

Table 4.7: List of Python packages/libraries used

Package/ library	Description	Link to documentation
Scikit-learn	An open-source machine learning library with support for supervised and unsupervised learning, model fitting, model selection and model evaluation.	https://scikit-learn.org/stable/
Matplotlib	An extensive library for creating interactive visualizations.	https://matplotlib.org/
Pandas	An open-source data analysis and manipulation tool.	https://pandas.pydata.org/
Numpy	A library with support for large, multi-dimensional arrays and matrices, and with an extensive collection of high-level mathematical functions.	https://numpy.org/
Joblib	Provides tools for lightweight pipelining in Python.	https://joblib.readthedocs.io/en/latest/



	Elevation (m)	Slope (°)	Aspect (°)	Surface Roughness	Topographic Position Index	Terrain Ruggedness Index	Terrain Surface Texture	Vector Ruggedness Measure
0	71.5352	2.997300	27.37760	4.367070	0.193008	3.96415	0.658470	0.000031
1	67.7092	0.128548	43.56470	0.315361	0.054283	0.31610	0.682026	0.000014
2	103.8560	2.311170	344.92300	2.940090	-0.010086	2.98073	0.180239	0.000007
3	68.8859	2.046760	7.80959	2.418470	0.106773	2.59480	0.947493	0.000062
4	95.4447	1.731710	347.01800	2.259900	0.196526	2.32884	0.208586	0.000053

Figure 4.19: Accessing the data using the read_csv() function from the Pandas library

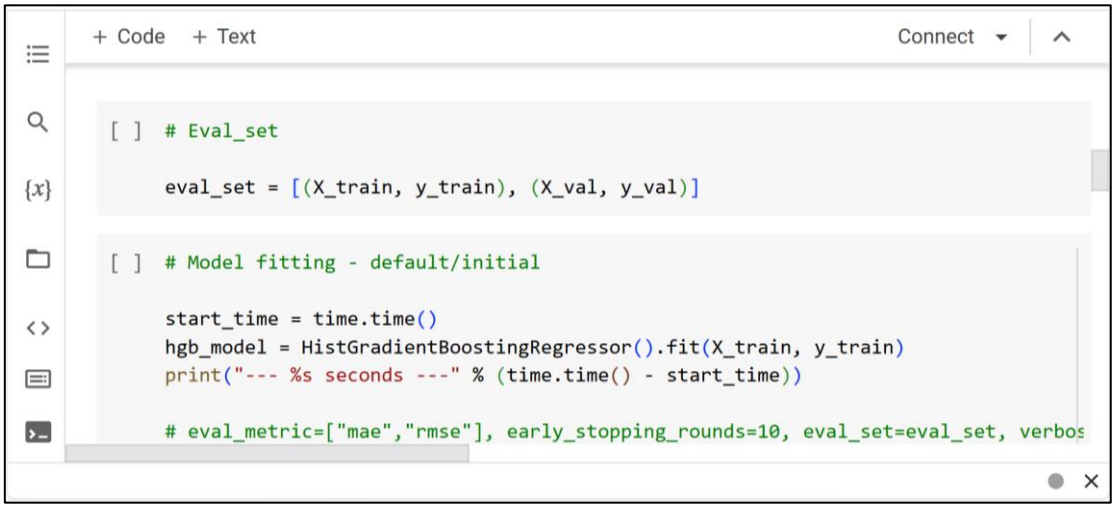


Figure 4.20: Data is passed into the histogram-based GBM regressor for training

4.8.5 Regression evaluation metrics

The Scikit-Learn metrics module incorporates several functions for measuring regression performance (Figure 4.21). The following metrics were adopted: root mean squared error (RMSE), mean absolute error (MAE), and median absolute error (MdAE). If $\hat{y}_i$ is the predicted elevation error of the i-th sample and y_i is the corresponding true value of the elevation error for a total of n samples, the estimated metrics are defined as (Chai & Draxler, 2014; Scikit-Learn, 2023a):

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \epsilon_i^2 \quad (4.11)$$

$$\mathrm{MSE}(y, \hat{y}) = \frac{1}{n}\sum_{i=0}^{n-1}(y_i - \hat{y}_i)^2 \quad (4.12)$$

$$\mathrm{MAE}(y, \hat{y}) = \frac{1}{n}\sum_{i=0}^{n-1}|y_i - \hat{y}_i| \quad (4.13)$$

$$\mathrm{MdAE}(y, \hat{y}) = \mathrm{median}(|y_1 - \hat{y}_1|, \ldots, |y_n - \hat{y}_n|) \quad (4.14)$$

The RMSE was derived by setting the 'squared' argument in the MSE function within the Scikit-learn machine learning library to 'False'.

Using the Python time module (Python, 2023), the execution times of the models for training and testing were also computed.

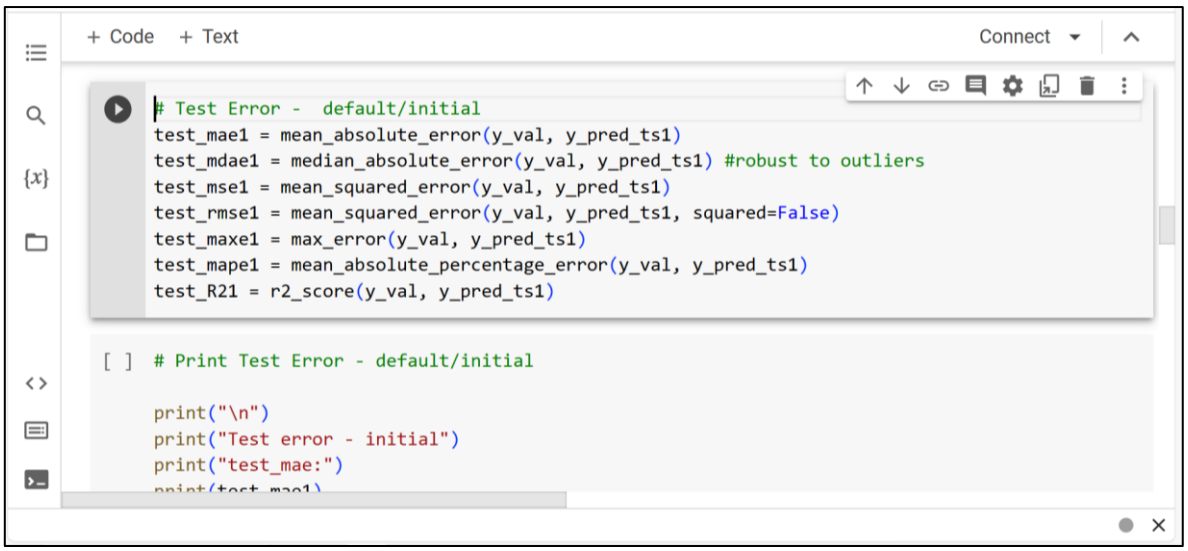


Figure 4.21: Regression error metrics

4.8.6 Explainability of tree-based ensembles

Explainability is crucial step for quantifying desirable properties of the evaluated models such as fairness, robustness, and causality (e.g., Adler & Painsky, 2022; Doshi-Velez & Kim, 2017).

It enables the identification of interdependencies within the system (Linardatos et al., 2020). Two techniques are adopted to address the model explainability: feature importance and partial dependence plots (PDPs), e.g., Figure 4.22. PDPs are model-agnostic plots that are used for describing a predictor's contribution to the fitted model (Molnar et al., 2022; Szepannek & Lübke, 2022).

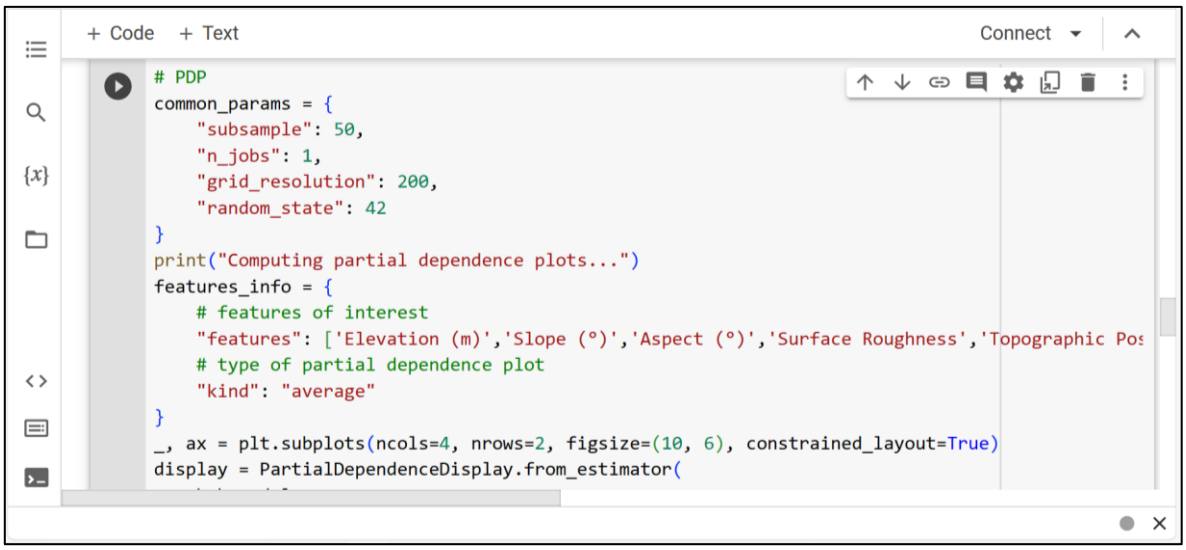


Figure 4.22: Partial dependence plots with the Scikit-learn machine learning library

4.9 Development of a GBDT-based DEM Enhancement Framework

Following the previous evaluation, this section focuses on the development of a comprehensive framework for DEM enhancement using three gradient boosted trees - XGBoost, LightGBM and CatBoost. The implementation was done in the Google Colab python development environment with an nvidia-smi GPU hardware accelerator.

4.9.1 Methodological framework for the GBDT-based DEM Enhancement framework

The selected sites for this evaluation are as shown in Figure 4.23. Figure 4.24 presents the methodological framework of the comprehensive DEM enhancement scheme. The workflow incorporates Bayesian optimisation for tuning the model hyperparameters.

4.9.2 Selected sites and data distribution

The distribution of training/test data and implementation sites for AW3D and Copernicus DEMs are visualized in Figure 4.23, and presented in Table 4.8 and Table 4.9 respectively. The 4th and 5th model implementation sites are introduced with letter designations corresponding to the landscapes, i.e., A – urban/industrial, B – agricultural, C – mountain, D – peninsula, and E – grassland/shrubland. This site arrangement is also maintained for the statistical-based

correction explained in section 4.11 Views of the landscapes are presented in Figure 4.25 - Figure 4.29.

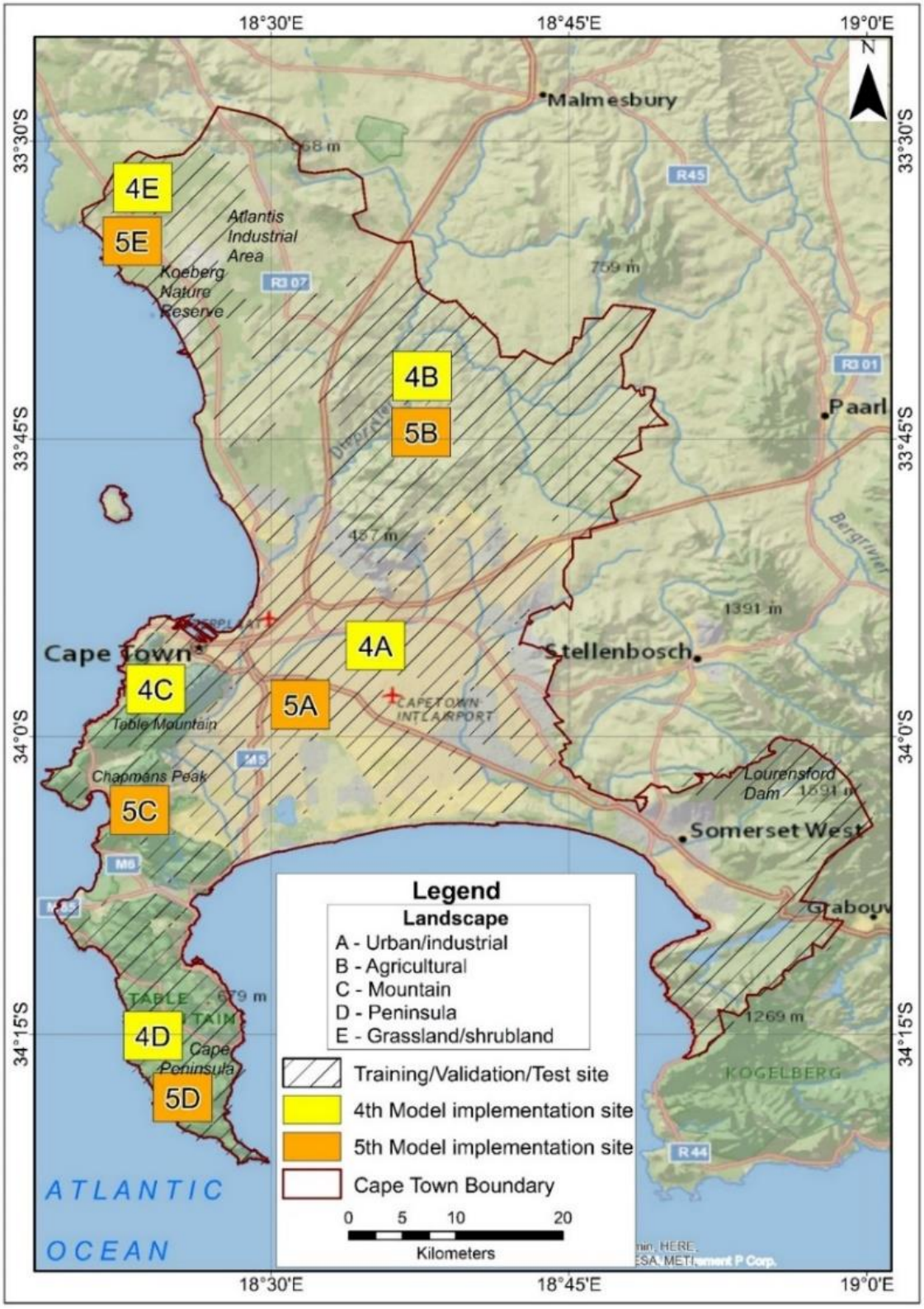


Figure 4.23: Selected sites and data distribution for the GBDT-based enhancement

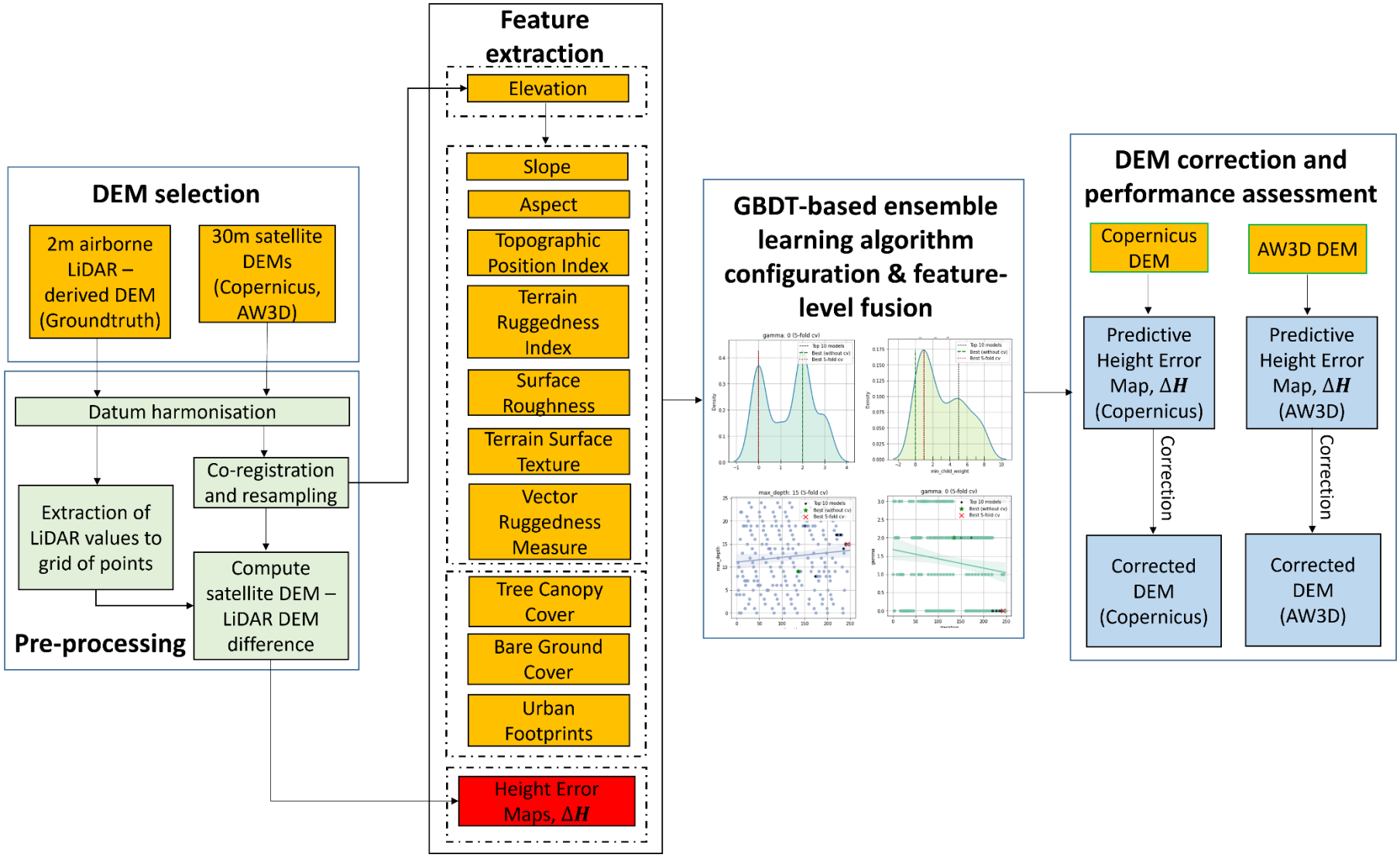


Figure 4.24: Workflow diagram of the comprehensive DEM enhancement scheme

Table 4.8: Distribution of data for GBDT-based enhancement – AW3D DEM

Landscape	Training/ validation/ test site		Model implementation site 4		Model implementation site 5	
	No of points	Area (km^2)	No of points	Area (km^2)	No of points	Area (km^2)
Urban/ industrial	572374	532.9	23041	20.7	22988	20.7
Agricultural	549555	509.0	23095	20.7	23082	20.7
Mountain	245929	280.6	23028	20.7	22722	20.7
Peninsula	129317	135.2	20916	20.7	20683	20.7
Grassland/ shrubland	125333	121.9	22658	20.7	22021	20.7

Table 4.9: Distribution of data for GBDT-based enhancement – Copernicus DEM

Landscape	Training/ validation/ test site		Model implementation site 4		Model implementation site 5	
	No of points	Area (km^2)	No of points	Area (km^2)	No of points	Area (km^2)
Urban/ industrial	573377	532.9	23041	20.7	22988	20.7
Agricultural land	549211	509.0	23095	20.7	23082	20.7
Mountain	251539	280.6	23028	20.7	22722	20.7
Peninsula	130859	135.2	20916	20.7	20683	20.7
Grassland/ shrubland	125378	121.9	22658	20.7	22021	20.7

Figure 4.25: View of urban/industrial landscapes selected for the GBDT-based DEM enhancement - training/test site (top), 4th implementation site (middle), 5th implementation site (bottom)

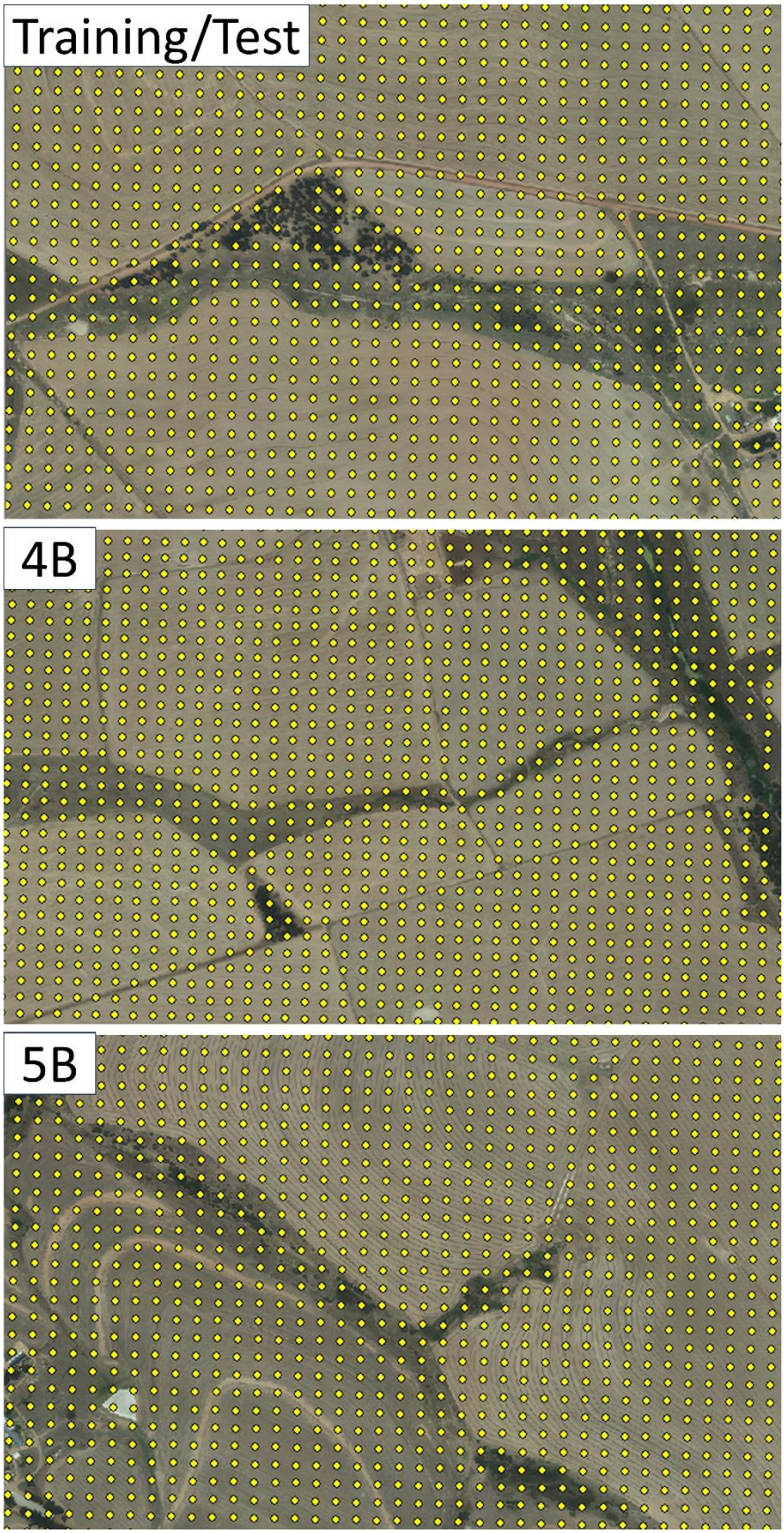


Figure 4.26: View of agricultural landscapes selected for the GBDT-based DEM enhancement - training/test site (top), 4th implementation site (middle), 5th implementation site (bottom). Training/test data points are superimposed on the image

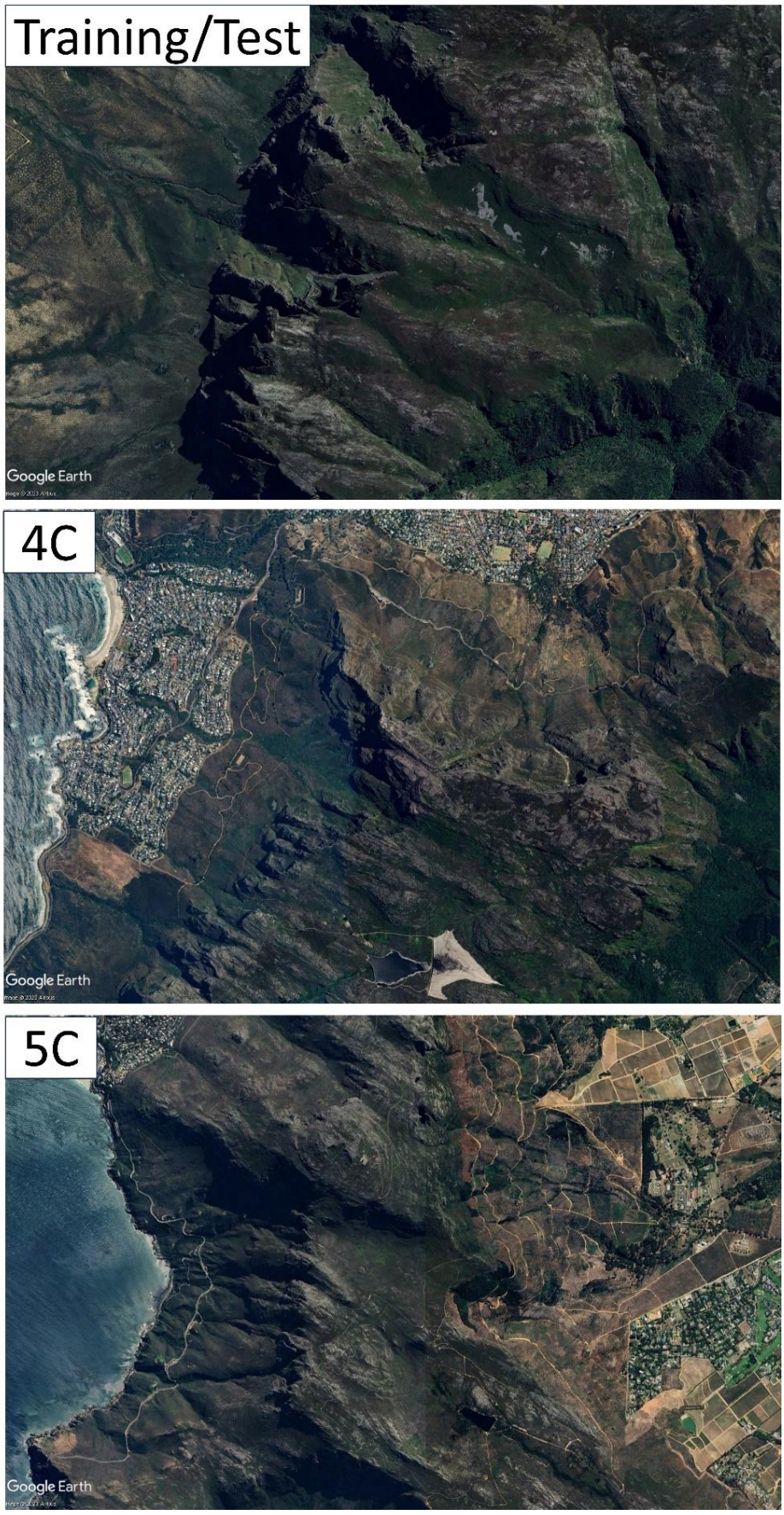


Figure 4.27: View of mountain landscapes selected for the GBDT-based DEM enhancement - training/test site (top), 4th implementation site (middle), 5th implementation site (bottom)

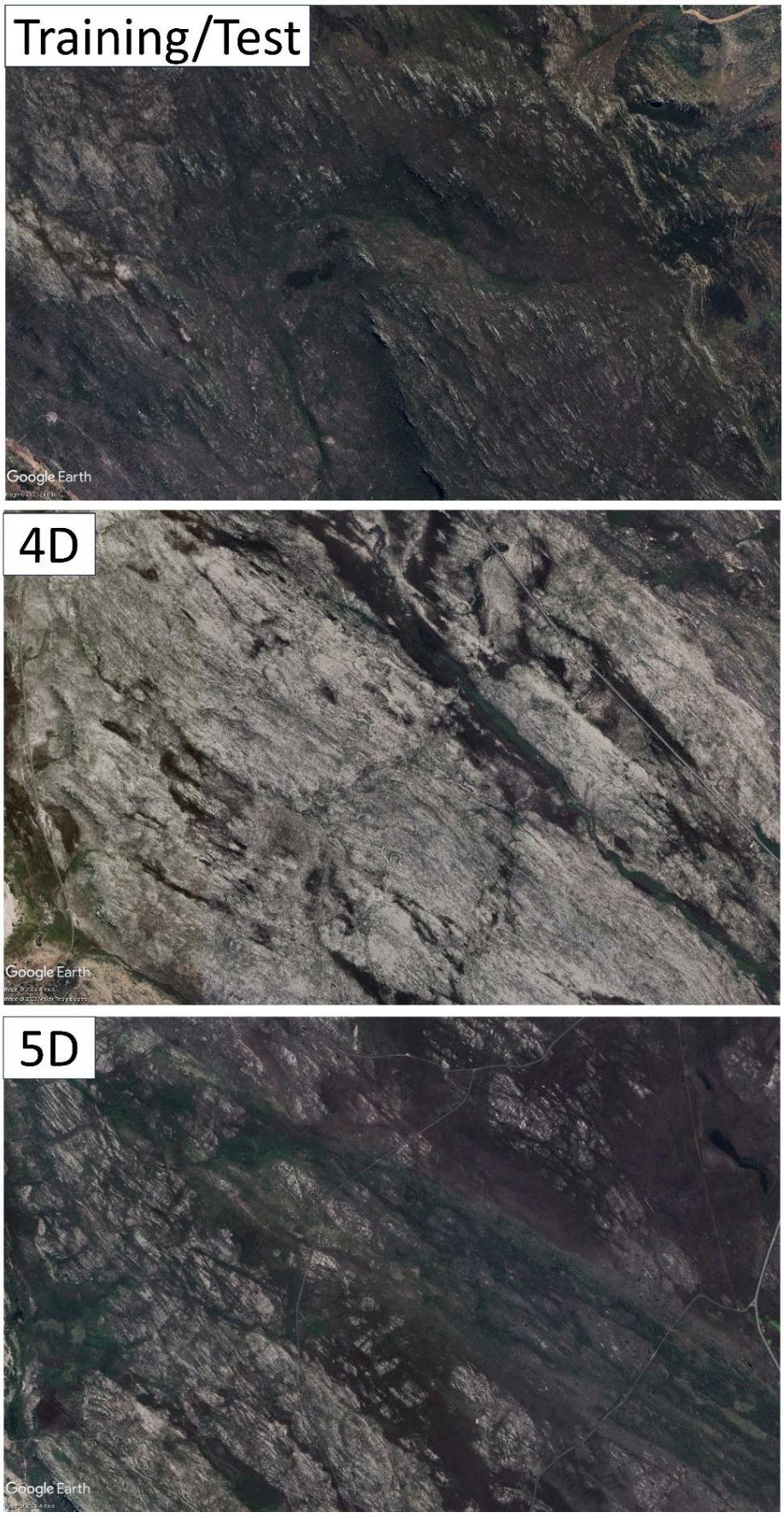


Figure 4.28: View of peninsula landscapes selected for the GBDT-based DEM enhancement - training/test site (top), 4th implementation site (middle), 5th implementation site (bottom)

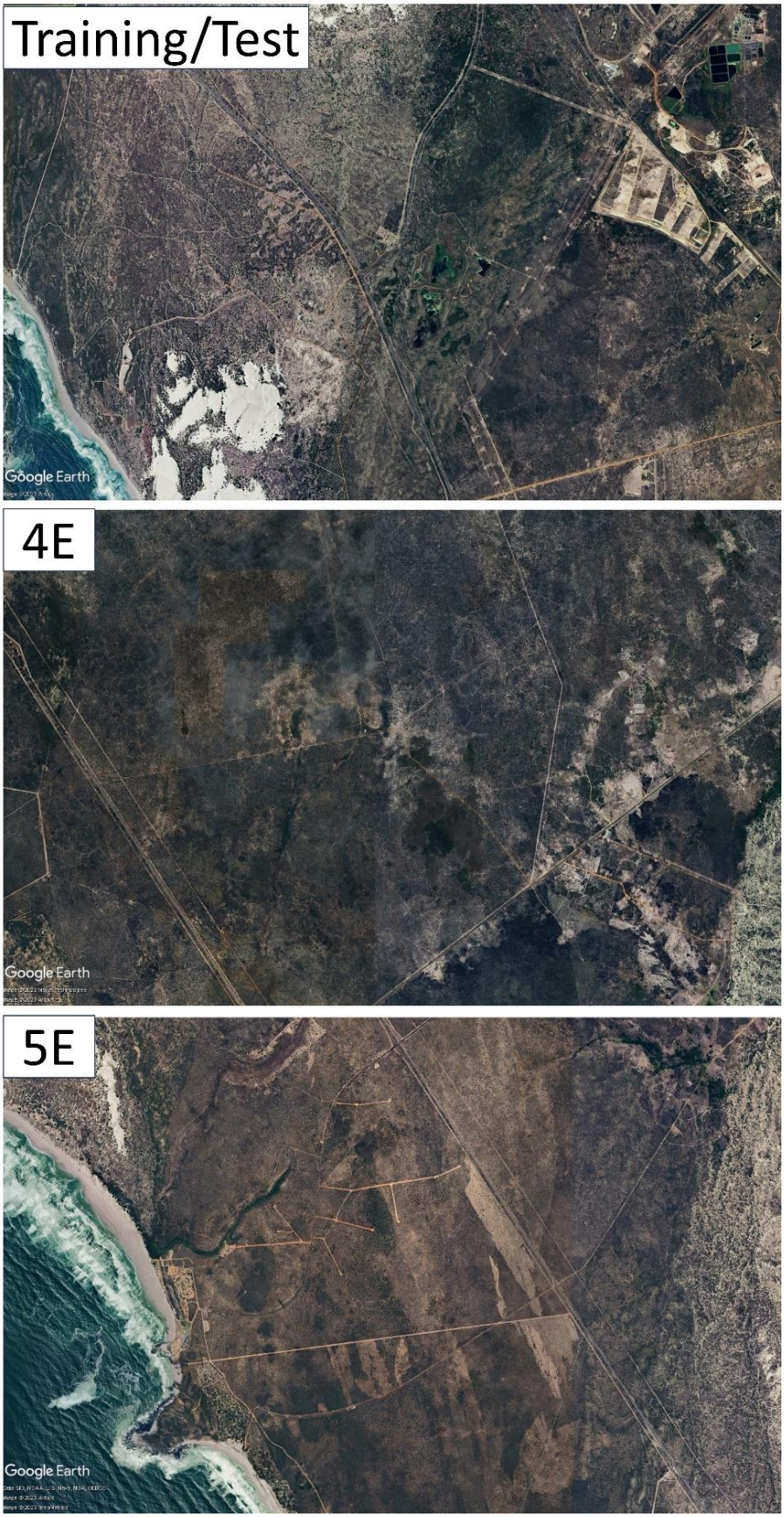


Figure 4.29: View of grassland/shrubland landscapes selected for the GBDT-based DEM enhancement - training/test site (top), 4th implementation site (middle), 5th implementation site (bottom)

4.9.3 Model implementation and Bayesian optimisation

The model implementation is addressed using both default and tuned (or optimised) hyperparameters. The first round of model implementation addressed both DEMs (AW3D and Copernicus) using default hyperparameters, while the second round of hyperparameter optimisation focused on Copernicus DEM. Some additional packages were used and their descriptions are shown in Table 4.10. Model hyperparameters control the learning process, and the chosen values influence the speed of training, accuracy and overfitting. Manual tuning of these hyperparameters is a game of trial and error, and very difficult to reproduce. Automated methods have been developed such as grid search and random search. However, the grid search is not scalable for higher dimensional problems, and random search settles for a local optima, and is reportedly incapable of deriving global optima (Victoria & Maragatham, 2021). Bayesian optimisation (BO) overcomes these constraints and was adopted in this research for identifying and selecting the most promising hyperparameter values. It is a good choice for this research because it has outperformed other state-of-the-art global optimisation algorithms on several challenging optimisation benchmark functions (Snoek et al., 2012).

Figure 4.30 presents the workflow for Bayesian optimisation of the model hyperparameters in this research.

Table 4.10: Additional packages for the GBDT- based DEM enhancement phase

Name of Python package/library	Description	Link to documentation
SHAP	Implements SHapley Additive exPlanations (SHAP) for explaining the output of machine learning models.	https://shap.readthedocs.io/en/latest/index.html
PDPbox	Partial dependence plot toolbox	https://pdpbox.readthedocs.io/en/latest/
BayesianOptimization	Implementation of Bayesian optimisation	https://github.com/bayesian-optimization/BayesianOptimization

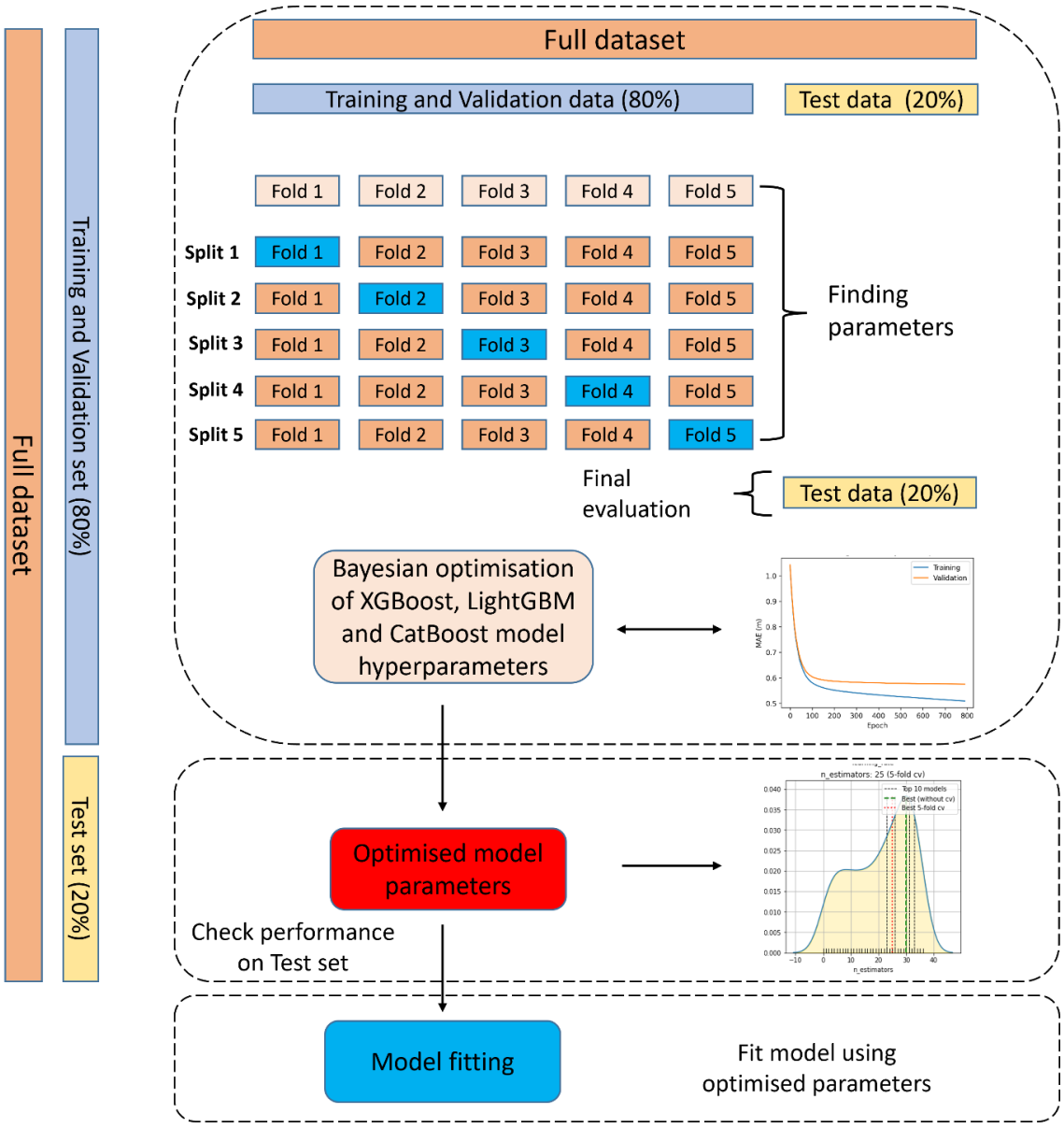


Figure 4.30: Workflow for Bayesian optimisation (Adapted from Taskesen, 2022, para. 1)

The Bayesian-optimization Python package (https://github.com/fmfn/BayesianOptimization) and other relevant modules were installed in the Google Colab environment, using the 'pip' Python package manager. Subsequently, the optimisation for XGBoost, LightGBM and CatBoost was meticulously implemented following this sequence:

Step 1 - Define objective function

The objective function was defined (e.g., Figure 4.31). This function takes hyperparameters as input and generates a score as output. The following general parameters were specified within the objective function:

1. **Early stopping rounds (early_stopping_rounds)**: This parameter helps to control overfitting and improve model performance by telling an algorithm when to stop if there is no further improvement in the evaluation metric. The value for early_stopping_rounds was set to 10 for all three models.
2. **K-fold cross-validation:** Cross validation is superior to the traditional machine learning data split mechanism. Five folds were specified for the K-fold cross validation where K represents the number of folds for splitting the data (i.e., 5-fold cross validation).
3. **Number of boosting rounds (**num_boost_round): The number of boosting rounds was set to 1000 for all three models.
4. **Evaluation metric** (eval_metric/metric): RMSE was adopted as the common evaluation metric.
5. **Boosting type/booster:**
   i. **XGBoost:** The 'gbtree' booster parameter was adopted.
   ii. **LightGBM:** The traditional gradient boosting decision tree, 'gbdt' was adopted.
   iii. **CatBoost:** The classic 'Plain' gradient boosting scheme was adopted.

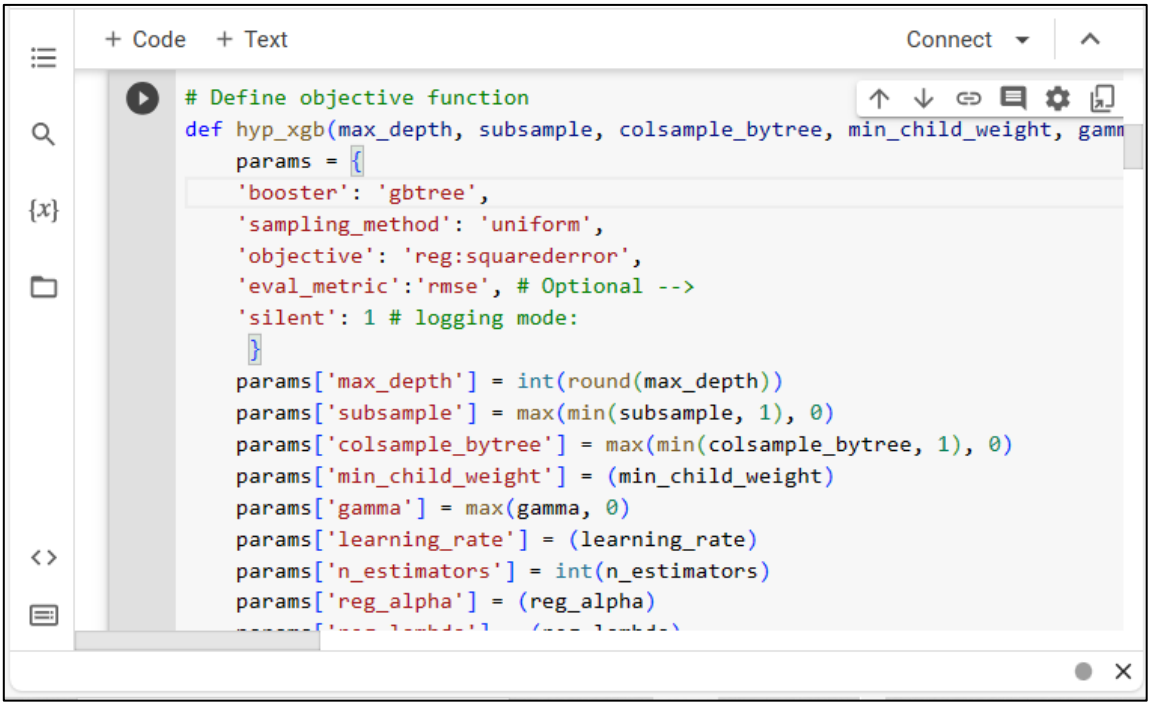


Figure 4.31: Defining the objective function for Bayesian optimisation of XGBoost

Step 2 - Defining the search space

Hyperparameter optimisation is seriously complicated by the wide array of potentially relevant hyperparameters available for tree-based models. Thus, the selection of the model

hyperparameters and the search space is not a trivial process. The initial selection considered the relevance of the hyperparameters based on their influence on model accuracy, speed of training and overfitting. Subsequently, an extensive series of tests were conducted whereby several hyperparameters were removed/replaced, and the effect on model performance was evaluated. This process also involved adjusting the search space and observing the effects. The search space was kept as wide as possible, within the limits of available computational resources (e.g., Figure 4.32). Finally, nine hyperparameters each were selected per model.

Brief descriptions of the hyperparameters, the default hyperparameter values and the unique search space defined for the Bayesian optimisation in this research are shown in Table 4.11 - Table 4.16 respectively. Further descriptions of the hyperparameters are available in the respective online documentations:

a. XGBoost - https://xgboost.readthedocs.io/en/stable/parameter.html
b. LightGBM - https://lightgbm.readthedocs.io/en/stable/Parameters.html
c. CatBoost - https://catboost.ai/en/docs/

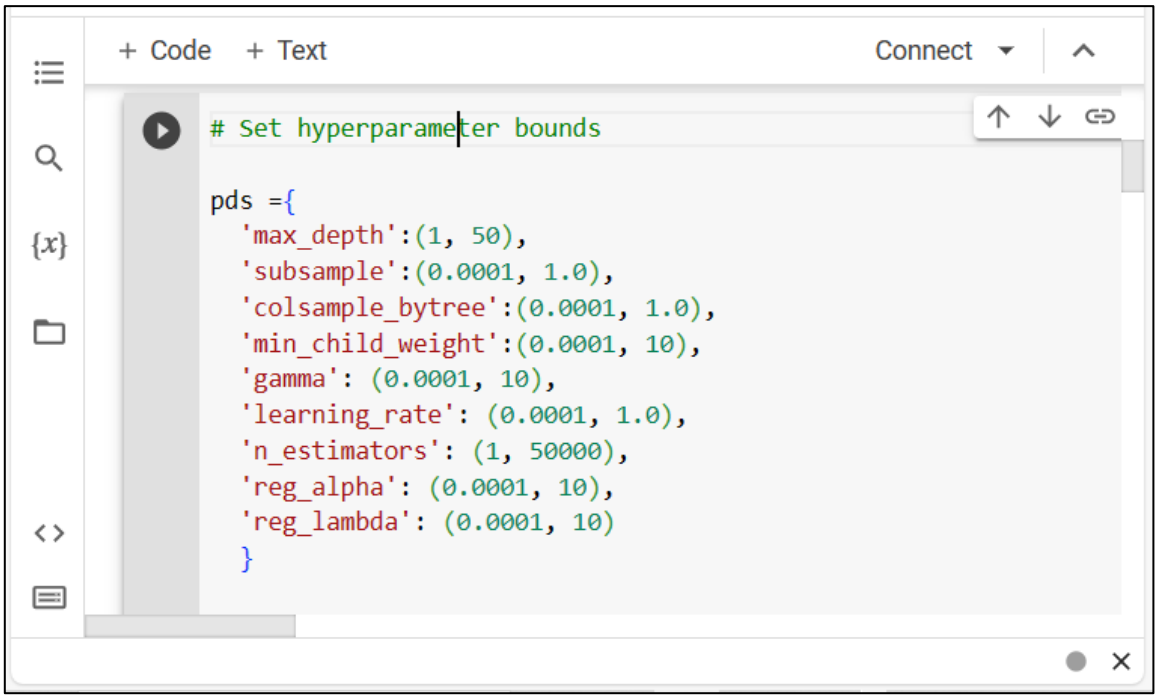


Figure 4.32: Defining the hyperparameter search space for Bayesian optimisation of XGBoost

Table 4.11: Description of model hyperparameters for Bayesian optimisation of XGBoost

Hyperparameter	Brief description	Value type
n_estimators	The number of decision trees to be boosted	Integer
learning_rate	Learning rate. It shrinks feature weights in each boosting step.	Float
max_depth	The maximum depth of tree. It limits the depth of tree growth.	Integer
reg_alpha	L1 regularization parameter. Increasing the value of L1 makes the model more conservative.	Float
reg_lambda	L2 regularization parameter. Increasing the value of L2 makes the model more conservative.	Float
min_child_weight	The minimum sum of instance weight required in a child.	Float
colsample_bytree	The subsample ratio of columns during the construction of each tree	Float
subsample	It represents subsample ratio of training instances.	Float
gamma	The minimum loss reduction to make a further partition on a leaf node. Used for tree pruning.	Float

Source: (Alam, 2021; Brownlee, 2016; XGBoost Developers, 2022)

Table 4.12: Comparison of the default hyperparameters versus the uniquely defined hyperparameter search space – XGBoost

Hyperparameter	Default value	Search space defined for Bayesian optimisation
n_estimators	100	(1, 50000)
learning_rate	0.3	(0.0001, 1)
max_depth	6	(1, 50)
reg_alpha	0	(0.0001, 10)
reg_lambda	1	(0.0001, 10)
min_child_weight	1	(0.0001, 10)
colsample_bytree	1	(0.0001, 1.0)
subsample	1	(0.0001, 1.0)
gamma	0	(0.0001, 10)

Table 4.13: Description of model hyperparameters for Bayesian optimisation of LightGBM

Hyperparameter	Brief description	Value type
num_iterations	The number of boosting rounds to be performed. This can also be regarded as the 'number of trees'	Integer
learning_rate	Learning rate. The rate at which LightGBM model weights are updated	Double
max_depth	The maximum depth of tree	Integer
lambda_l1	L1 regularization parameter. Increasing the value of L1 makes the model more conservative.	Double
lambda_l2	L2 regularization parameter. Increasing the value of L2 makes the model more conservative.	Double
min_child_weight	The minimum sum of instance weight required in a child	Double
feature_fraction	A subset of features selected on each iteration (tree)	Double
min_gain_to_split	The minimal gain to perform a split.	Double
bagging_fraction	Similar to feature_fraction, but randomly selects a subset of data without resampling	Double

Source: (Bahmani, 2023; Microsoft Corporation, 2023)

Table 4.14: Comparison of the default hyperparameters versus the uniquely defined hyperparameter search space – LightGBM

Hyperparameter	Default value	Search space defined for Bayesian optimisation
num_iterations	100	(1, 50000)
learning_rate	0.1	(0.0001, 1)
max_depth	-1	(1, 50)
lambda_l1	0	(0.0001, 10)
lambda_l2	0	(0.0001, 10)
min_child_weight	0.001	(0.0001, 1)
feature_fraction	1	(0.0001, 1)
min_gain_to_split	0	(0.0001, 1)
bagging_fraction	1	(0.0001, 1)

Table 4.15: Description of model hyperparameters for Bayesian optimisation of CatBoost

Hyperparameter	Brief description	Value type
iterations	The maximum number of trees that can be built.	Integer
learning_rate	Learning rate. It is used for reducing the gradient step.	Float
depth	The depth of the tree	Integer
l2_leaf_reg	Coefficient for the L2 regularization term of the cost function.	Float
colsample_bylevel	The random subspace method. The percentage of features to be used at each split selection, when features are randomly selected.	Float
random_strength	The amount of randomness for scoring splits when the tree structure is selected.	Float
leaf_estimation_iterations	The number of newton/gradient steps done in every tree when calculating leaf values	Integer
bagging_temperature	The bagging temperature defines the settings of the Bayesian bootstrap.	Float
border_count	Border count. This is the number of splits for numerical features.	Integer

Source: (CatBoost, 2023)

Table 4.16: Comparison of the default hyperparameters versus the uniquely defined hyperparameter search space – CatBoost

Hyperparameter	Default value	Search space defined for Bayesian optimisation
iterations	1000	(1, 50000)
learning_rate	0.03	(0.0001, 1)
depth	6	(1,16)
l2_leaf_reg	3.0	(0.0001, 10)
colsample_bylevel	None (set to 1)	(0.0001, 1)
random_strength	1	(0.0001, 10)
leaf_estimation_iterations	None (Depends on the training objective)	(1, 10)
bagging_temperature	1	(0.0001, 10)
border_count	CPU: 254; GPU: 32 or 128	(1, 255)

Step 3 - Surrogate model

A surrogate probability model of the objective function was defined.

Step 4 – Run optimisation

The Bayesian optimisation function was called to maximise the objective output (e.g., Figure 4.33). The following parameters were specified:

a. **No. of random explorations (init_points):** 10 steps of random exploration were specified. This enabled the diversification of the exploration space.
b. **Number of iterations (n_iter):** 40 - 50 steps of Bayesian optimisation were specified. Increasing the number of steps increases the likelihood of finding a good maximum. For Copernicus DEM in the urban/industrial, mountain, peninsula and grassland/shrubland landscapes, 50 iterations were executed for the Bayesian optimisation of all three models, while 40 iterations were executed for each model in the agricultural landscape.

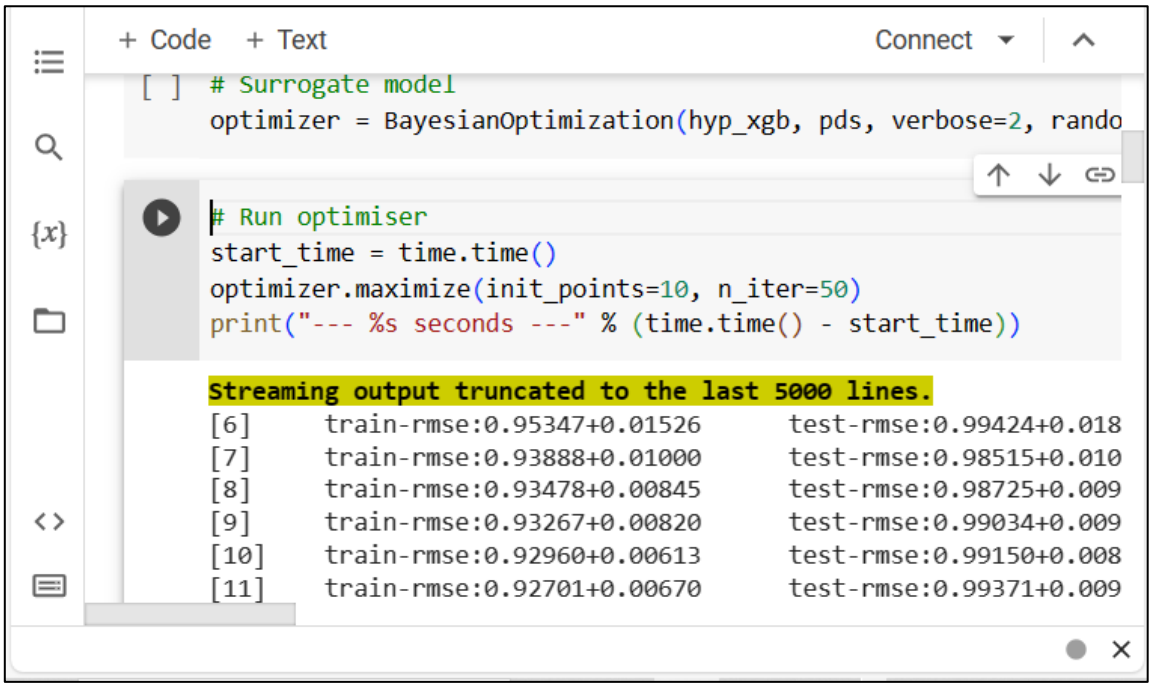


Figure 4.33: Defining the surrogate model and running the optimisation for XGBoost

The optimisation time ranged from several minutes to hours. Thereafter, the best combination of hyperparameters and target values was accessed using the property, ‘optimizer.max’ (Figure 4.34). The list of probed parameters and their target values were accessed using the property, ‘optimizer.res’ (Figure 4.35). Thereafter, the models were trained using the derived hyperparameter values (e.g., Figure 4.36).

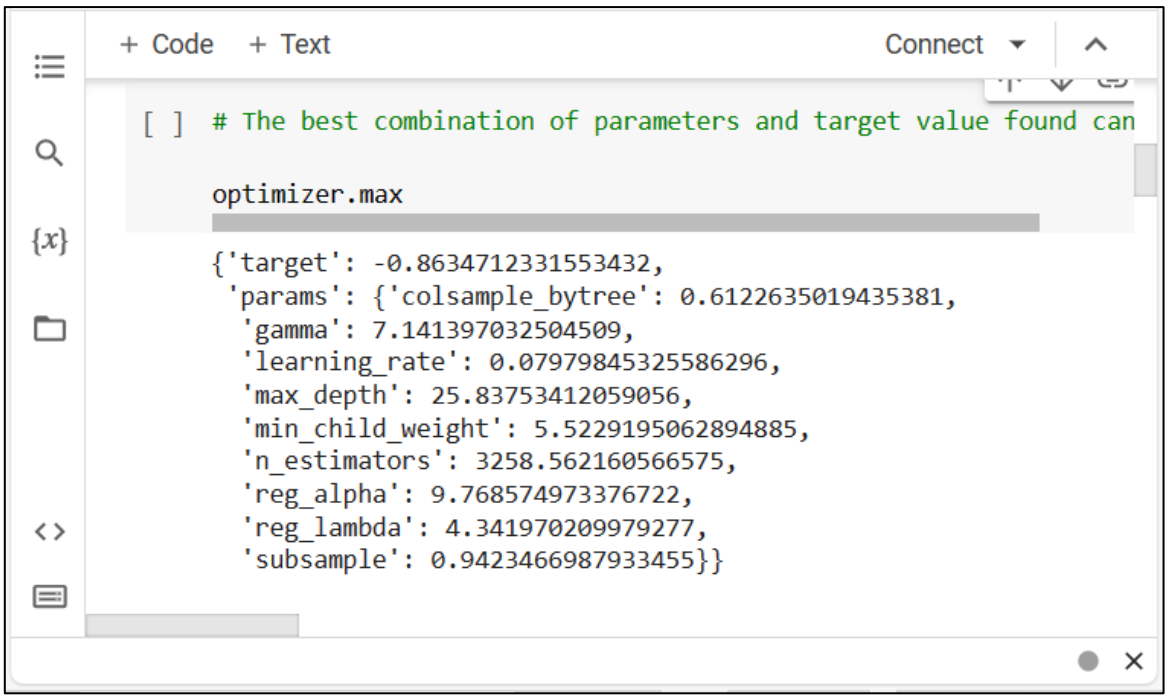


Figure 4.34: Generating the best combination of hyperparameters

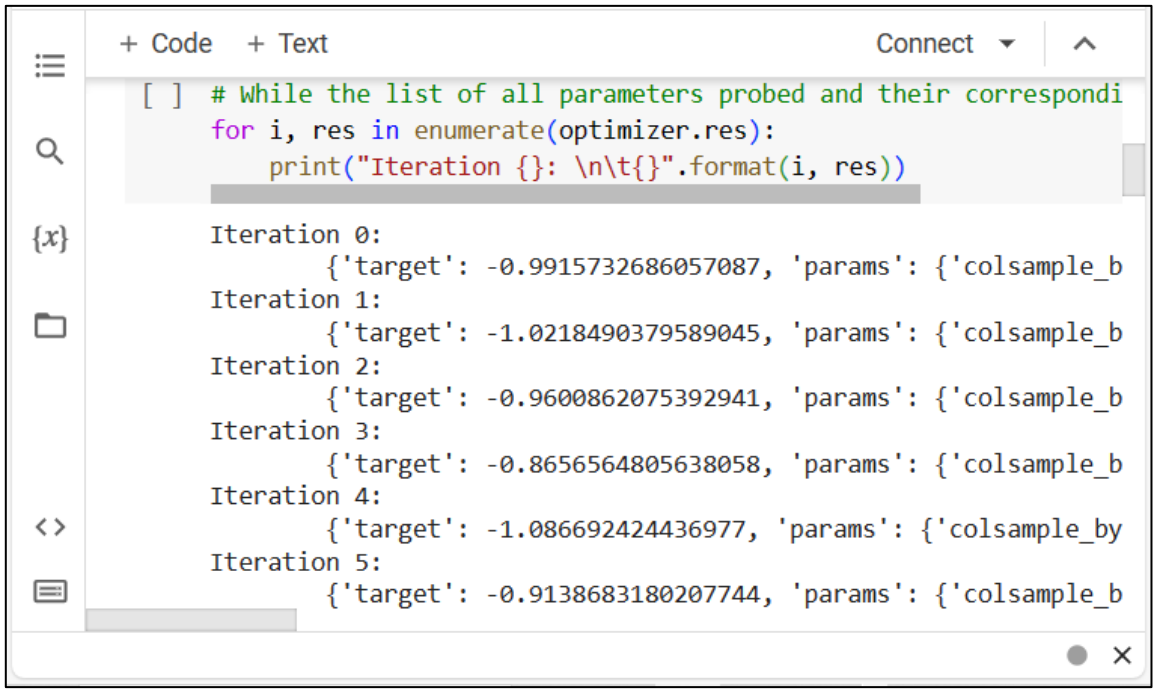


Figure 4.35: Displaying the list of probed hyperparameters for each iteration and corresponding target values

Figure 4.36: Training the XGBoost model using optimised hyperparameter values

4.9.4 Explainability of GBDT-based enhancement results

The explainability of the GBDT-based results was addressed with the use of shapley additive explanations (SHAP). Firstly, SHAP was installed using the pip Python package installer (Figure 4.37). To compute SHAP values, an Explainer object needs to be created and used for evaluating a sample or the full dataset (Trevisan, 2022). The SHAP explainer was fit to the three models (XGBoost, LightGBM and CatBoost) and the results were visualized using SHAP feature importance and beeswarm plots.

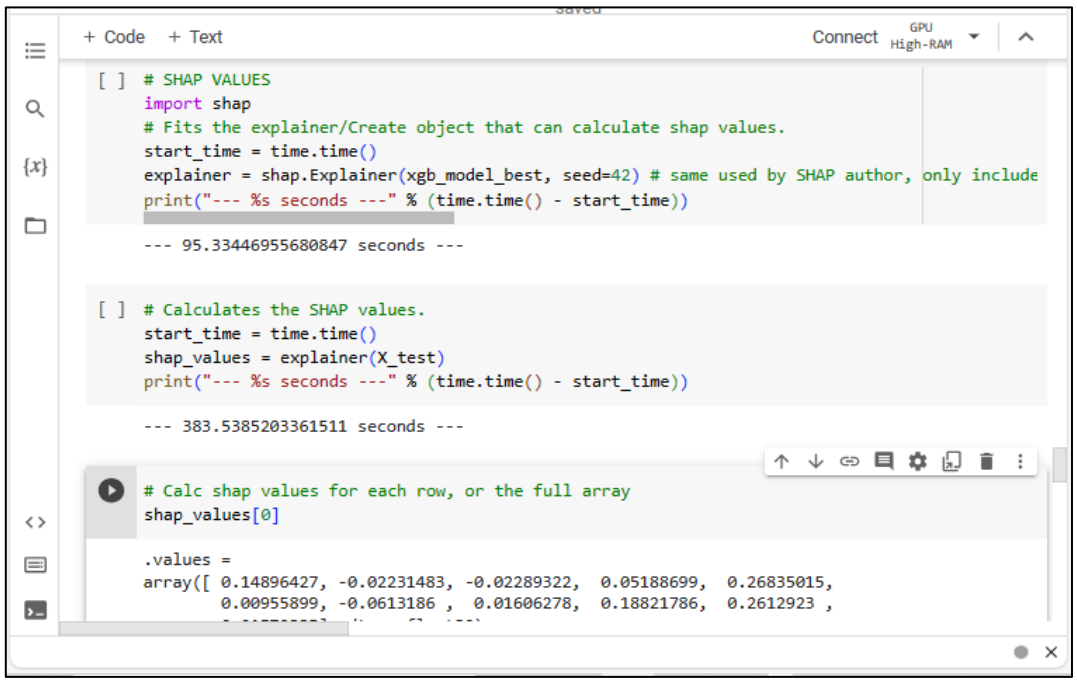


Figure 4.37: Fitting the explainer and calculating SHAP values

4.10 Performance Evaluation of GBDT-based DEM Enhancement Framework

4.10.1 Terrain visualisation and analysis

The potential of the enhanced DEMs for improved terrain visualisation and analysis was evaluated at selected sites through a qualitative comparison of 10 m interpolated (and smoothened) contours, 3D terrain surfaces and shaded relief maps from the LiDAR DEM, original Copernicus DEM, corrected Copernicus DEM and South Africa National DEM. The contours were interpolated and smoothened within the ArcGIS environment. The PAEK (Polynomial Approximation with Exponential Kernel) method was adopted for smoothening the contour lines to remove sharp angles in the lines for the purpose of improving their aesthetic appearance and cartographic quality (Figure 4.38).

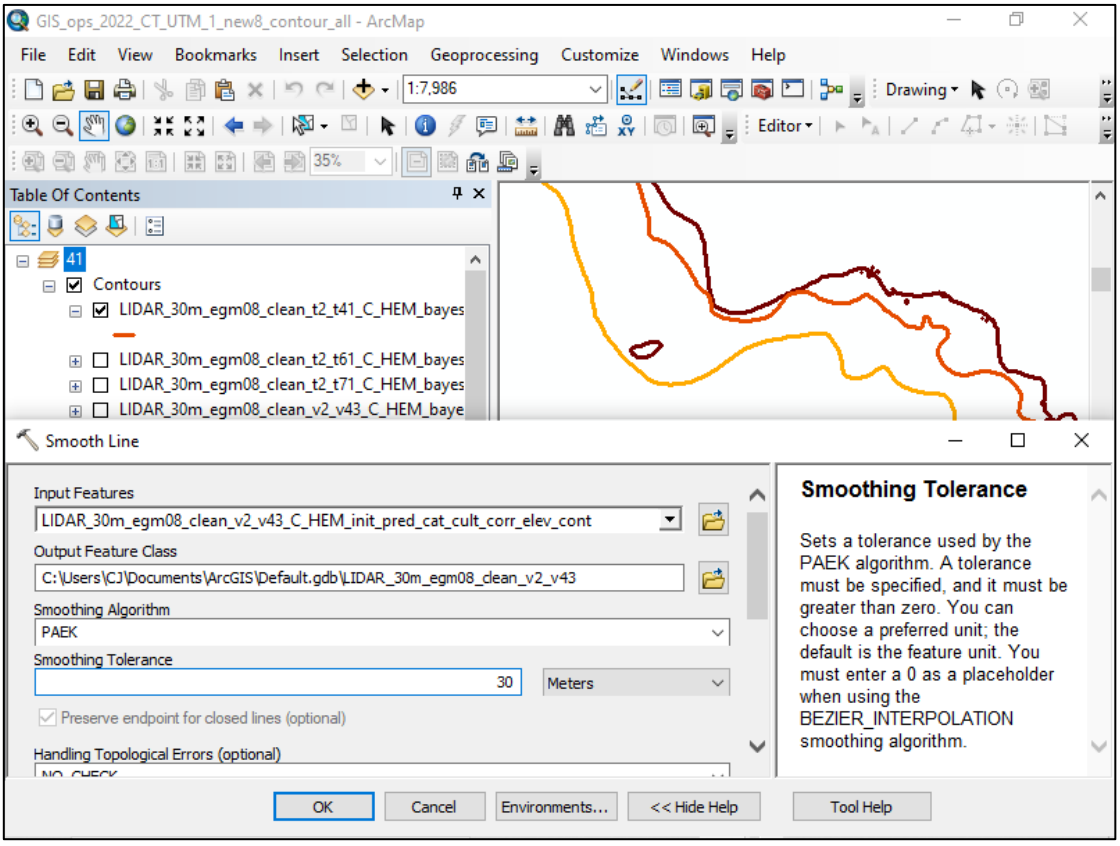
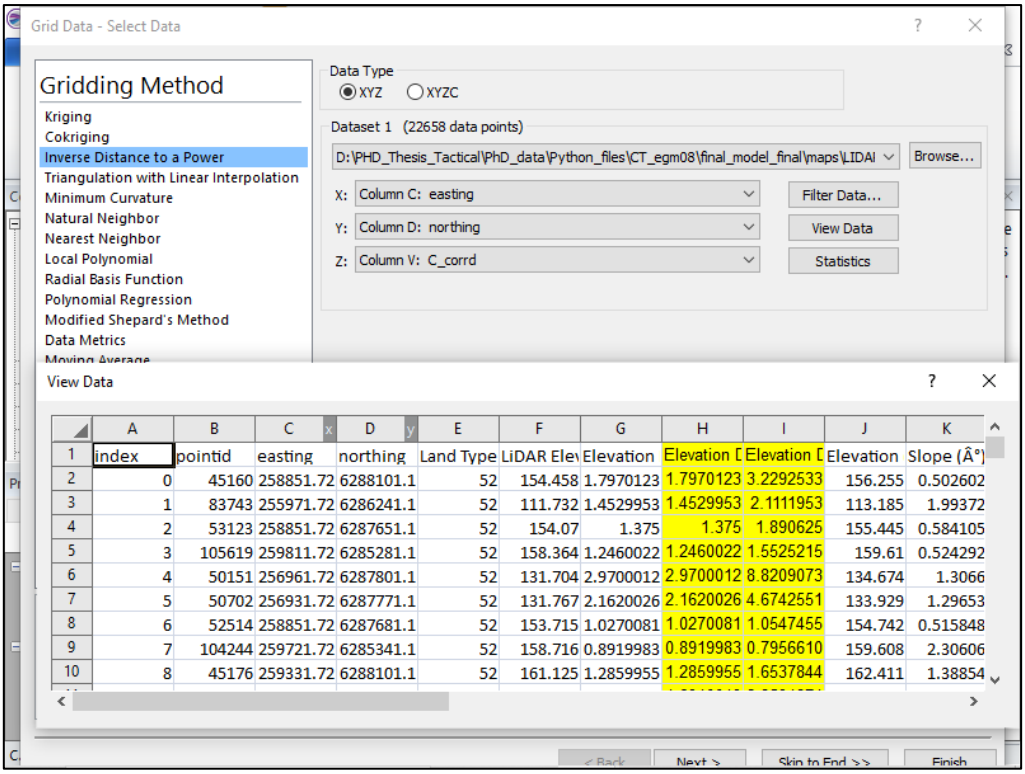


Figure 4.38: Smoothening the contours using the PAEK method

The next evaluation involved a comparison of 3D terrain surfaces and shaded relief maps generated within the Surfer software environment. Surfer is a 3D modelling software with extensive tools for terrain visualisation (Golden Software, 2023). The 3D surfaces and shaded relief maps were generated from gridded XYZ files which were interpolated from the elevation data points using the IDW method (Figure 4.39). The heights on the 3D surfaces correspond to the height value (or Z value) of the associated grid node (Figure 4.40).



	A	B	C	D	E	F	G	H	I	J	K
1	index	pointid	easting	northing	Land Type	LiDAR Ele	Elevation	Elevation [	Elevation [	Elevation	Slope (Â°)
2	0	45160	258851.72	6288101.1	52	154.458	1.7970123	1.7970123	3.2292533	156.255	0.502602
3	1	83743	255971.72	6286241.1	52	111.732	1.4529953	1.4529953	2.1111953	113.185	1.99372
4	2	53123	258851.72	6287651.1	52	154.07	1.375	1.375	1.890625	155.445	0.584105
5	3	105619	259811.72	6285281.1	52	158.364	1.2460022	1.2460022	1.5525215	159.61	0.524292
6	4	50151	256961.72	6287801.1	52	131.704	2.9700012	2.9700012	8.8209073	134.674	1.3066
7	5	50702	256931.72	6287771.1	52	131.767	2.1620026	2.1620026	4.6742551	133.929	1.29653
8	6	52514	258851.72	6287681.1	52	153.715	1.0270081	1.0270081	1.0547455	154.742	0.515848
9	7	104244	259721.72	6285341.1	52	158.716	0.8919983	0.8919983	0.7956610	159.608	2.30606
10	8	45176	259331.72	6288101.1	52	161.125	1.2859955	1.2859955	1.6537844	162.411	1.38854

Figure 4.39: Generating an XYZ grid on Surfer

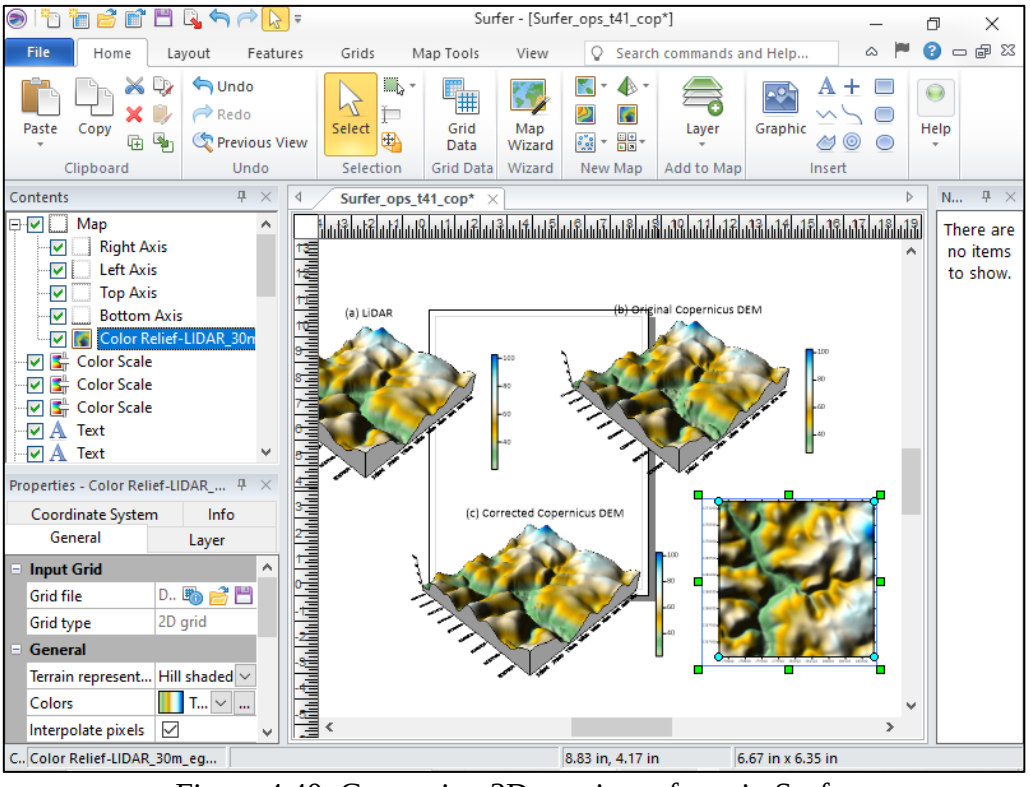


Figure 4.40: Generating 3D terrain surfaces in Surfer

4.10.2 Mapping of lineaments

Lineaments are defined as “simple or composite linear features of a surface, the parts of which are aligned in a rectilinear or slightly curvilinear relationship uniquely different from the

pattern of adjoining features and that apparently reflects a sub-surface phenomenon" (O'Leary & Friedman, 1978 in Tiren, 2010, p. 10). Depending on their evolution, three types of lineaments may be distinguished (Ahmadi & Pekkan, 2021; Echeverria et al., 2022; Epuh et al., 2020; Minár & Sládek, 2009; O'leary et al., 1976): (i) geological lineaments (ii) morpho-structural/topographic lineaments, and (iii) non-geological lineaments. Lineament extraction is a very important application of remote sensing in geological mapping (Echeverria et al., 2022). DEMs play a crucial role by facilitating lineament extraction in inaccessible and challenging mountainous areas (Shetty et al., 2021). Hill-shaded DEMs are an alternative to aerial photographs and satellite images for the extraction of lineaments (Henderson et al., 1996). Since DEM quality affects the resulting terrain derivatives, this research incorporates lineament extraction from the original and corrected DEMs in the performance assessment. The selected sites for lineament extraction are the Table Mountain landscapes. The lineament extraction followed a 2-step process:

Step 1: Derivation of shaded relief maps

Following the approach of Soliman & Han (2019), shaded relief images were derived from the DEMs in multiple illumination directions (0°, 45°, 90°, and 135°) at a sun elevation angle of 45°. The hillshading was carried out within SAGAGIS (Figure 4.41; Figure 4.42).

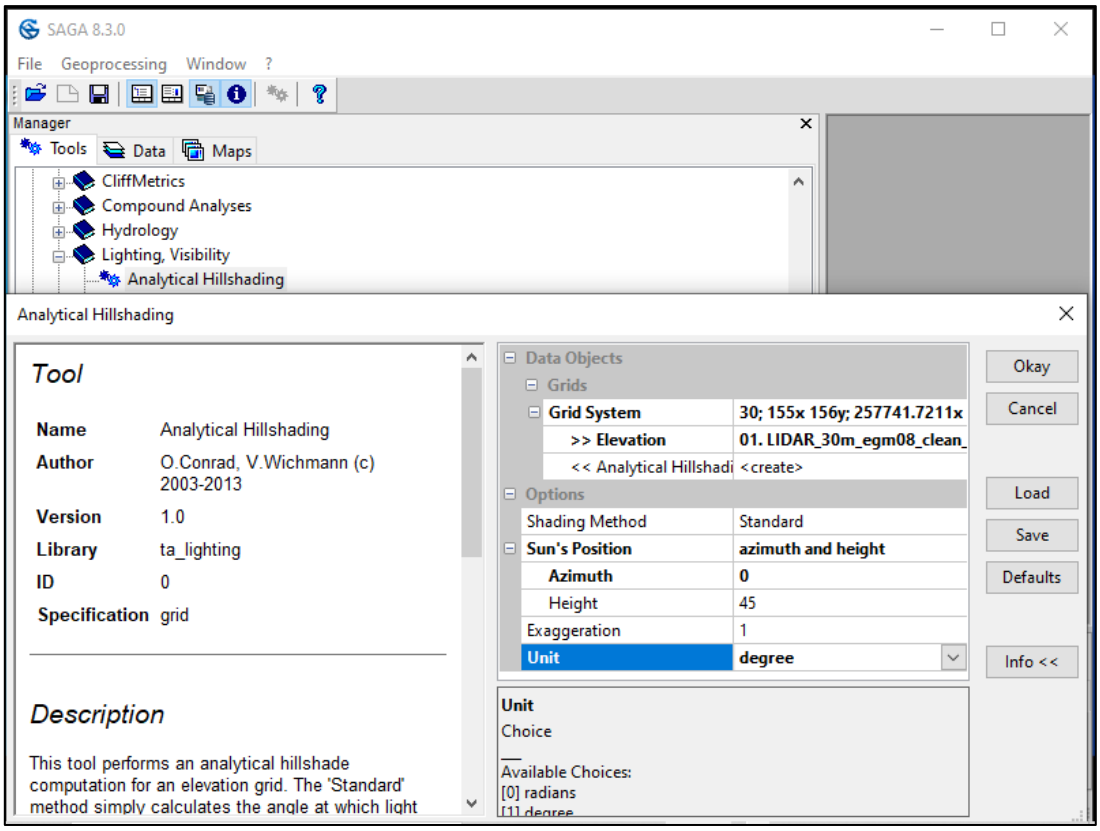


Figure 4.41: Analytical hillshading in SAGAGIS

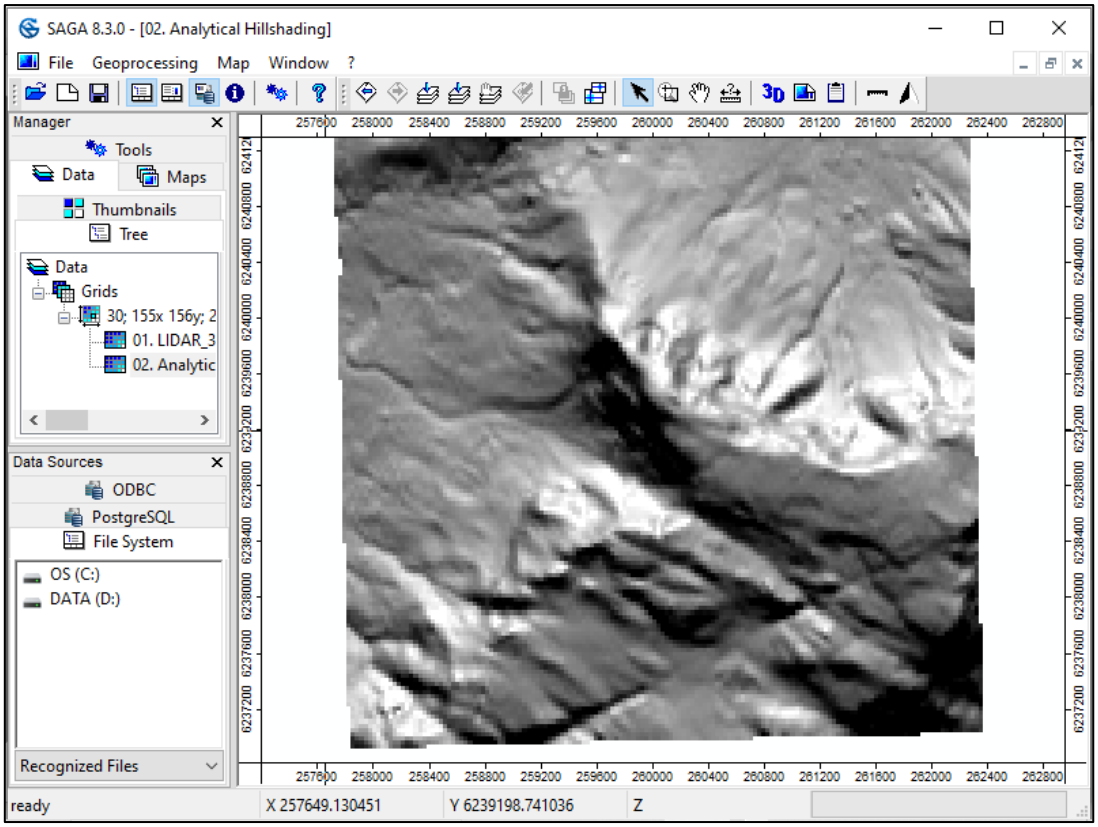


Figure 4.42: Output of hillshaded DEM in SAGAGIS

Step 2: Lineament extraction

The analytically shaded maps were imported into the CATALYST Professional software environment for edge-detection, thresholding and curve extraction with the LINE module (Figure 4.43 - Figure 4.44). CATALYST Professional is a geospatial software specializing in earth observation science, photogrammetry, and remote sensing with powerful image processing and analysis capabilities (PCI - Geomatics, 2022). The Line module enabled the extraction of lineaments from the analytically shaded DEMs and its sensitivity is dependent on the six parameters shown in Table 4.17. The explanation of these parameters is well documented (e.g., Echeverria et al., 2022). The extracted lineaments were thereafter converted to ArcGIS shapefiles. Within ArcGIS, the lineaments were split at vertices, and the following geometry parameters were computer for each line:

- Length of lineament
- X-coordinate (Easting) of lineament start point
- X-coordinate (Easting) of lineament endpoint
- Y-coordinate (Northing) of lineament start point
- Y-coordinate (Northing) of lineament endpoint

Thereafter, the lineaments were converted to DXF (drawing exchange format). The DXF files were imported into Rockworks software (Figure 4.45) for the preparation of rose diagrams (Figure 4.46). Rose diagrams enable the visualisation and analysis of the lineament orientation (i.e., the circular distribution of the directional lineaments). Rockworks is a comprehensive software by RockWare Inc. for creating geological models, 2D/3D maps, and general geology diagrams that are used in several industries (RockWare, 2023).

Table 4.17: Specified Line module parameters in the extraction of lineaments

Parameter	**Description**	**Specified value**
RADI	Filter radius (pixels)	5
GTHR	Edge gradient threshold	5
LTHR	Curve length threshold (pixels)	5
FTHR	Line fitting error threshold (pixels)	5
ATHR	Angular difference threshold (degrees)	5
DTHR	Linking distance threshold (pixels)	5

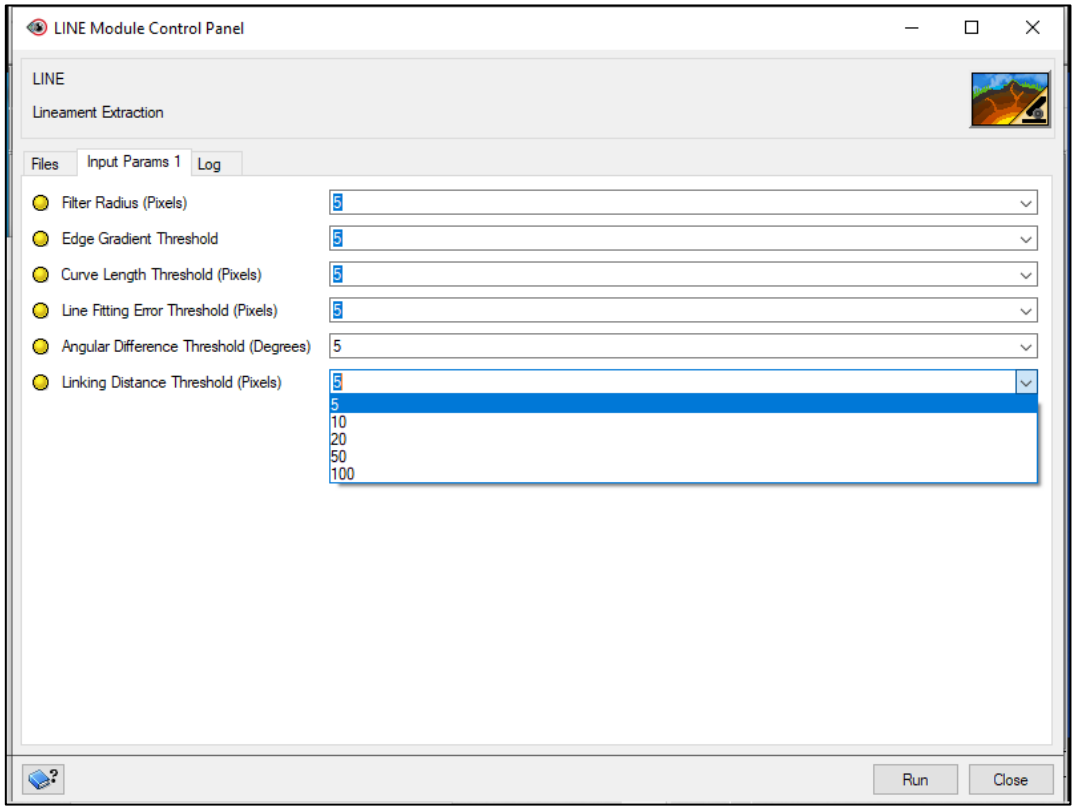


Figure 4.43: Specifying the parameters of the LINE module for lineament extraction in CATALYST Professional

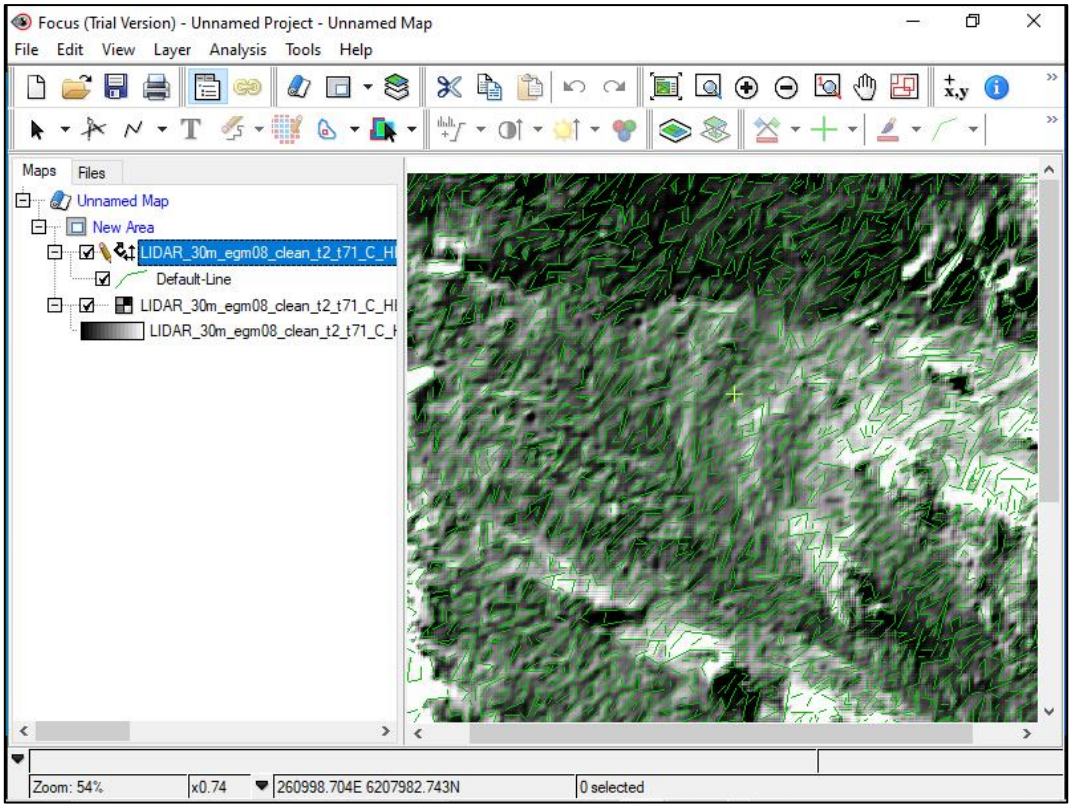
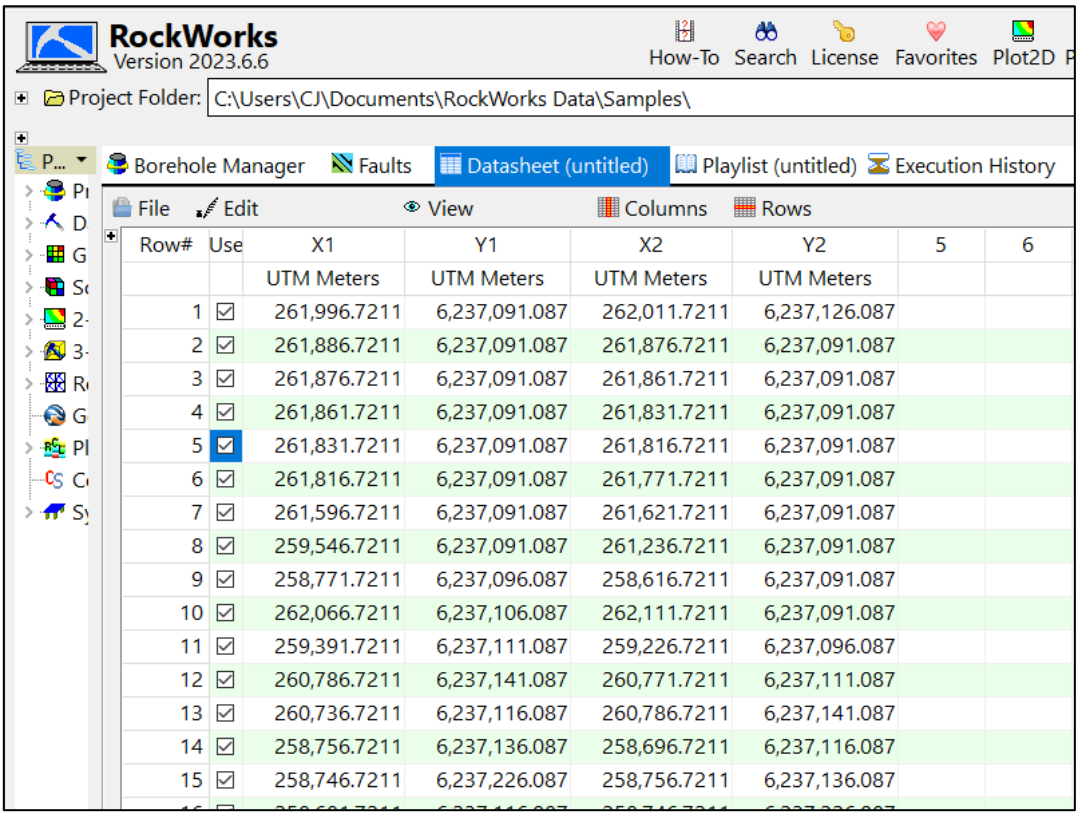


Figure 4.44: Extracted lineaments shown in the CATALYST Professional environment



Row#	Use	X1	Y1	X2	Y2	5	6
		UTM Meters	UTM Meters	UTM Meters	UTM Meters		
1	☑	261,996.7211	6,237,091.087	262,011.7211	6,237,126.087		
2	☑	261,886.7211	6,237,091.087	261,876.7211	6,237,091.087		
3	☑	261,876.7211	6,237,091.087	261,861.7211	6,237,091.087		
4	☑	261,861.7211	6,237,091.087	261,831.7211	6,237,091.087		
5	☑	261,831.7211	6,237,091.087	261,816.7211	6,237,091.087		
6	☑	261,816.7211	6,237,091.087	261,771.7211	6,237,091.087		
7	☑	261,596.7211	6,237,091.087	261,621.7211	6,237,091.087		
8	☑	259,546.7211	6,237,091.087	261,236.7211	6,237,091.087		
9	☑	258,771.7211	6,237,096.087	258,616.7211	6,237,091.087		
10	☑	262,066.7211	6,237,106.087	262,111.7211	6,237,091.087		
11	☑	259,391.7211	6,237,111.087	259,226.7211	6,237,096.087		
12	☑	260,786.7211	6,237,141.087	260,771.7211	6,237,111.087		
13	☑	260,736.7211	6,237,116.087	260,786.7211	6,237,141.087		
14	☑	258,756.7211	6,237,136.087	258,696.7211	6,237,116.087		
15	☑	258,746.7211	6,237,226.087	258,756.7211	6,237,136.087		

Figure 4.45: Importing points data into Rockworks

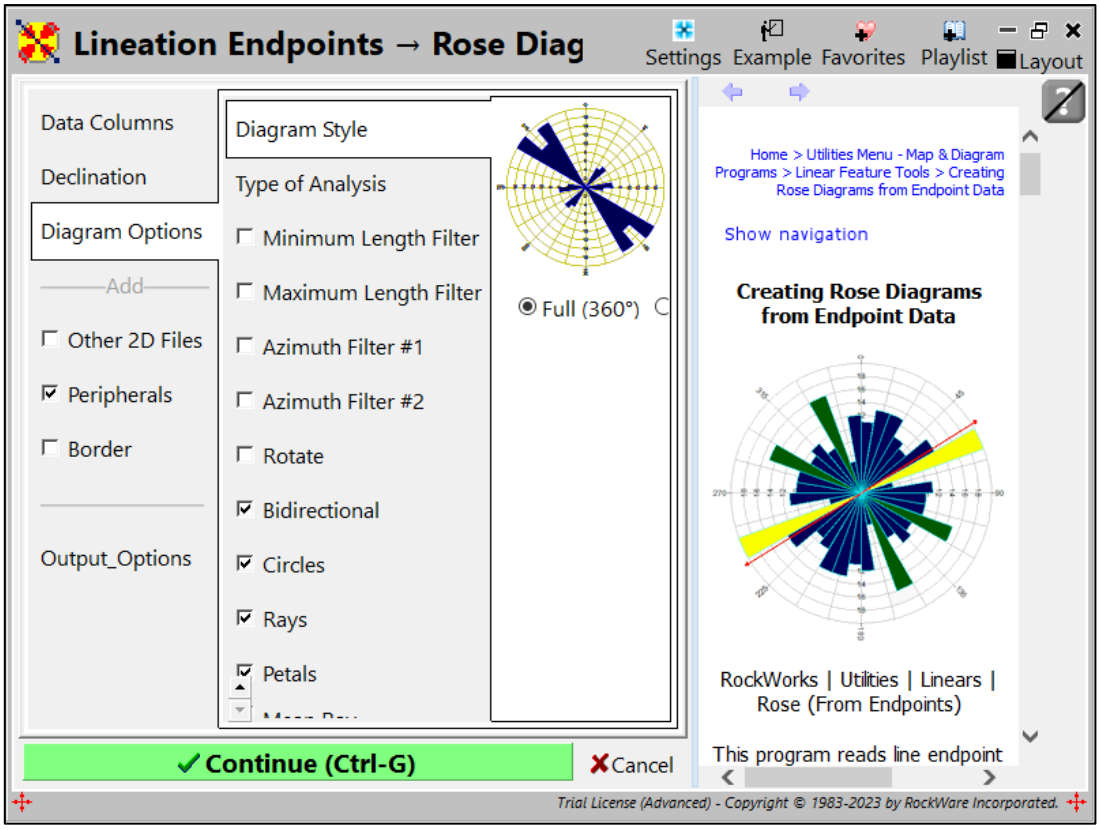


Figure 4.46: Parameter settings for generating rose diagrams in Rockworks

4.11 Comparison with Statistical-based Correction of DEMs

This section presents the methodology for the derivation of prediction models using multiple linear regression (MLR) for the correction of Copernicus and AW3D DEMs. The objective is to provide an alternative approach for comparing the results from machine learning. Unlike decision trees, MLR is not immune to multicollinearity. Thus, the first consideration was to screen the input variables for the presence of multi-collinearity. This was achieved using two tests: the correlation analysis and variance inflation factor (VIF).

4.11.1 Methodological framework for statistical-based correction of DEMs

The methodological framework for the multiple linear regression analysis is presented in Figure 4.47. Initially, the already co-registered DEMs, terrain and land cover parameters were selected. This was followed by multi-collinearity diagnostics to detect or flag collinearity in the terrain parameters. Thus, terrain parameters that are very highly inter-related were excluded in this stage. Two methods were adopted for the collinearity tests: Pearson's correlation analysis and variance inflation factor (VIF) tests. The final selected variables were then fed into the regression model.

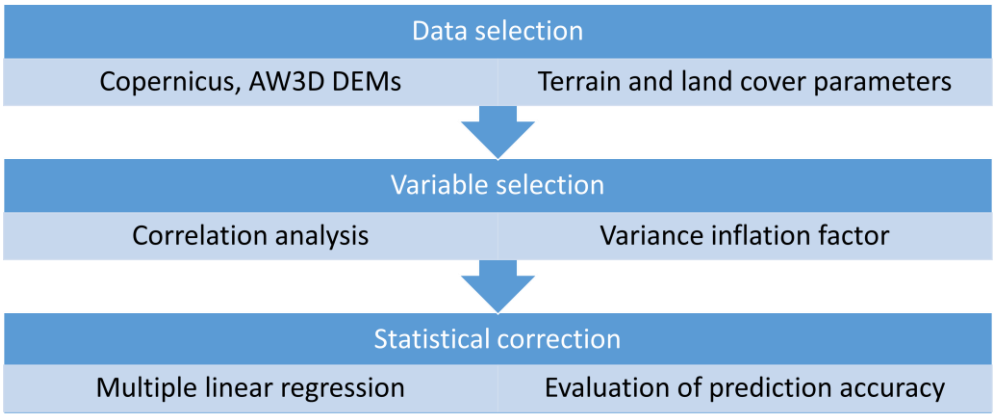


Figure 4.47: Workflow diagram for the regression analysis

4.11.2 Selected sites and data distribution

The selected sites for the regression analysis are shown in Figure 4.48, and the data distribution is as shown in Table 4.18 - Table 4.19 respectively.

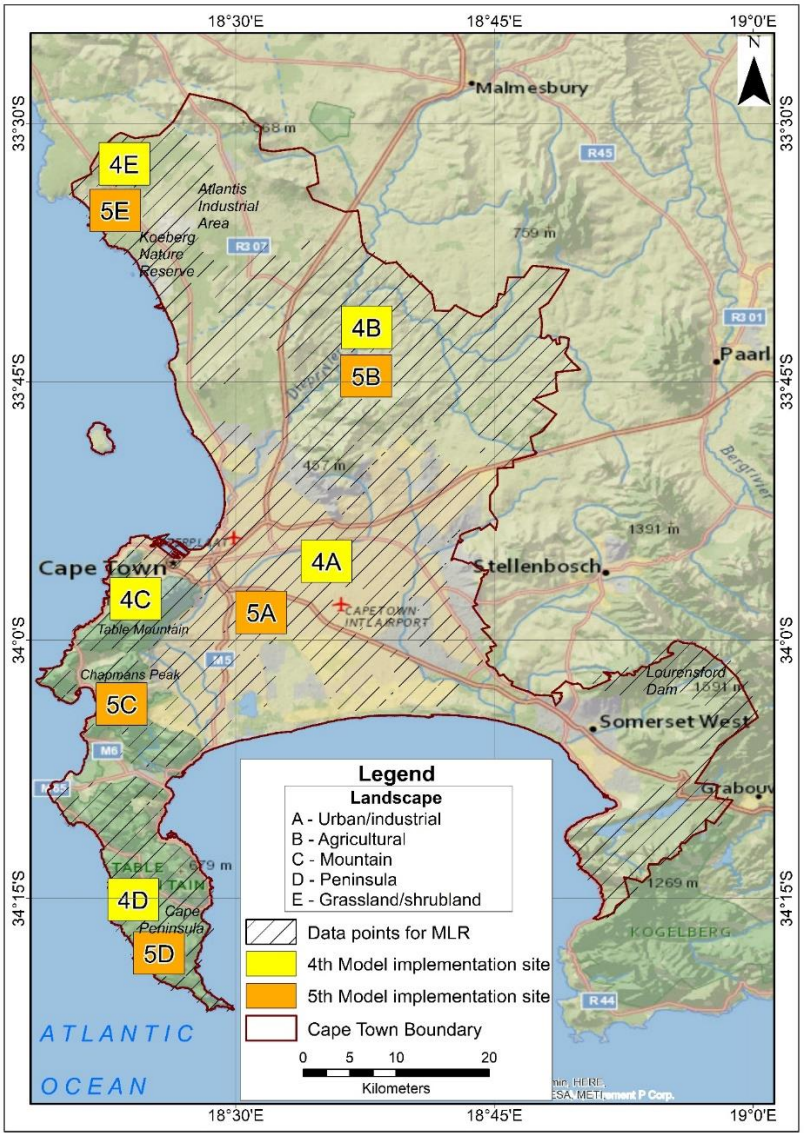


Figure 4.48: Selected sites and data distribution for regression analysis

Table 4.18: Distribution of data for statistical-based correction – AW3D DEM

Landscape	Training/ validation/ test site		Model implementation site 4		Model implementation site 5	
	No of points	Area (km^2)	No of points	Area (km^2)	No of points	Area (km^2)
Urban/ industrial	572374	532.9	23041	20.7	22988	20.7
Agricultural	549555	509.0	23095	20.7	23082	20.7
Mountain	245929	280.6	23028	20.7	22722	20.7
Peninsula	129317	135.2	20916	20.7	20683	20.7
Grassland/ shrubland	125333	121.9	22658	20.7	22021	20.7

Table 4.19: Distribution of data for statistical-based correction – Copernicus DEM

Landscape	Training/ validation/ test site		Model implementation site 4		Model implementation site 5	
	No of points	Area (km^2)	No of points	Area (km^2)	No of points	Area (km^2)
Urban/ industrial	573377	532.9	23041	20.7	22988	20.7
Agricultural land	549211	509.0	23095	20.7	23082	20.7
Mountain	251539	280.6	23028	20.7	22722	20.7
Peninsula	130859	135.2	20916	20.7	20683	20.7
Grassland/ shrubland	125378	121.9	22658	20.7	22021	20.7

4.11.3 Correlation analysis

Pearson's bivariate correlation analysis was carried out to flag any significant correlations between the input parameters. The Pearson product-moment correlation coefficient (r_{xy}) is given as (Alademomi et al., 2022; Nwilo et al., 2020):

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (4.15)$$

Where,

r_{xy} = correlation coefficient (ranging from +1.0 to -1.0)

x_i = values of the x-variable in the given sample

$\bar{x}$ = mean of the values of the x-variable

y_i = values of the y-variable in the given sample

$\bar{y}$ = mean of the values of the y-variable

The correlation analysis was implemented within SPSS software (Figure 4.49).

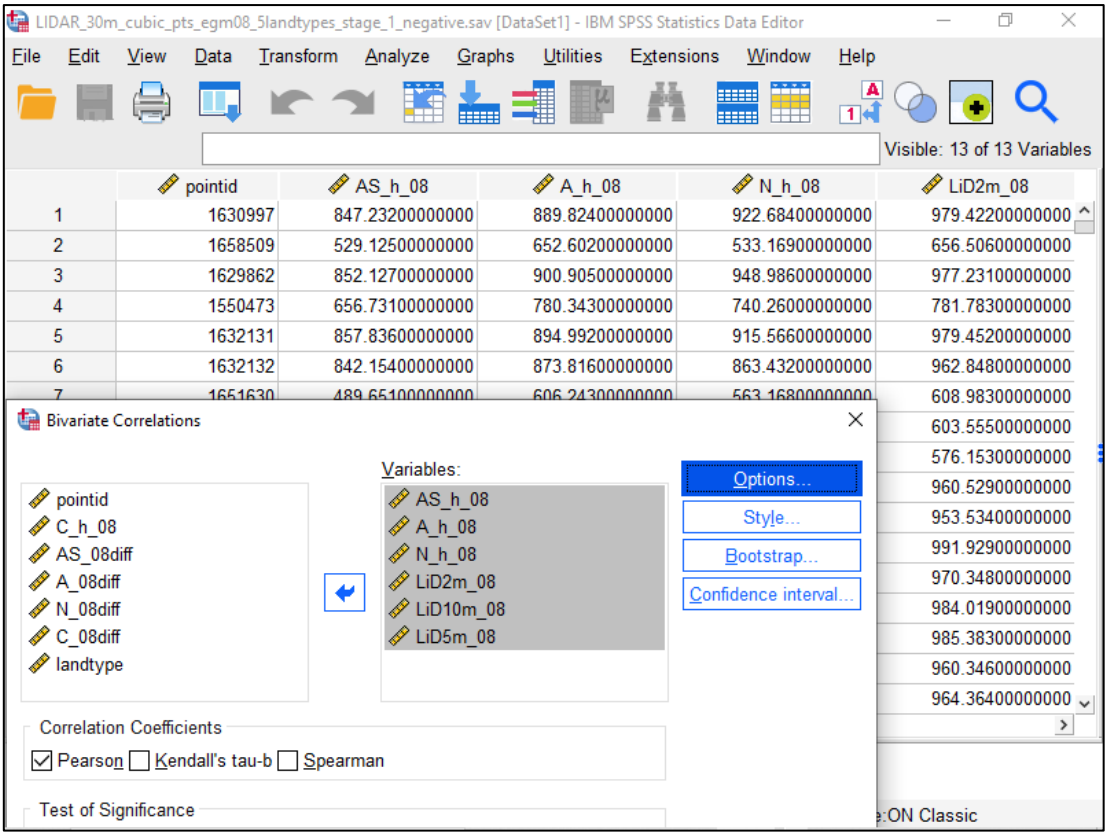


Figure 4.49: Correlation analysis with SPSS

4.11.4 Variance inflation factor

Variance Inflation Factor is a diagnostic measure for detecting multicollinearity, and is considered as a more robust diagnostic tool. The VIF can indicate the problematic coefficients that are impacted by collinearity (Ferré, 2009). Using a general rule of thumb, variables with VIF > 10 are usually eliminated. The VIF for the jth independent variable is given by Forthofer et al. (2007).

$$VIF_j = \frac{1}{1-R_j^2} \qquad (4.16)$$

Where,

R^2_j is the R^2 from the regression of the jth independent variable on the remaining independent variables. The VIF was implemented with Python scripting (Figure 4.50).

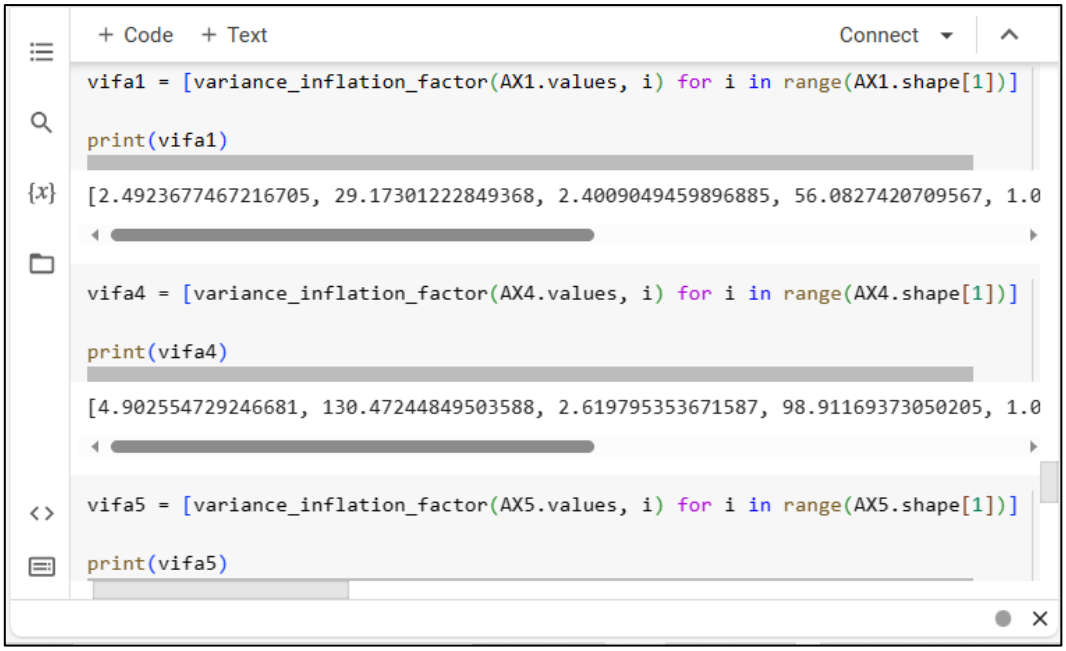


Figure 4.50: Python script for the variance inflation factor

4.11.5 Regression analysis for prediction of elevation error

It is assumed that the independent variables (i.e., elevation, slope, aspect, surface roughness, TPI, TRI, TST, VRM, urban footprint values, percentage tree cover and percentage bare ground) have an impact on the elevation/height error. Thus, using MLR, the elevation error can be expressed as a linear combination of the selected independent variables. An MLR model with p independent variables takes the form (Huang, 2023; Ong et al., 2015):

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_p x_{i,p} + \varepsilon_i \quad (4.17)$$

Where,

y_i = dependent variable, and $x_{i,1}$, $x_{i,2}$,…, $x_{i,p}$ are the independent variables

β is a p-dimensional parameter vector, and its elements are referred to as the regression coefficients. β_0 is the contant term.

ε_i = the error term.

The MLR was implemented within R, using the following syntax for the **lm()** function (RDocumentation, 2023) (Figure 4.51):

lm(formula, data, subset, weights, na.action,
method = "qr", model = TRUE, x = FALSE, y = FALSE, qr = TRUE,
singular.ok = TRUE, contrasts = NULL, offset, …)

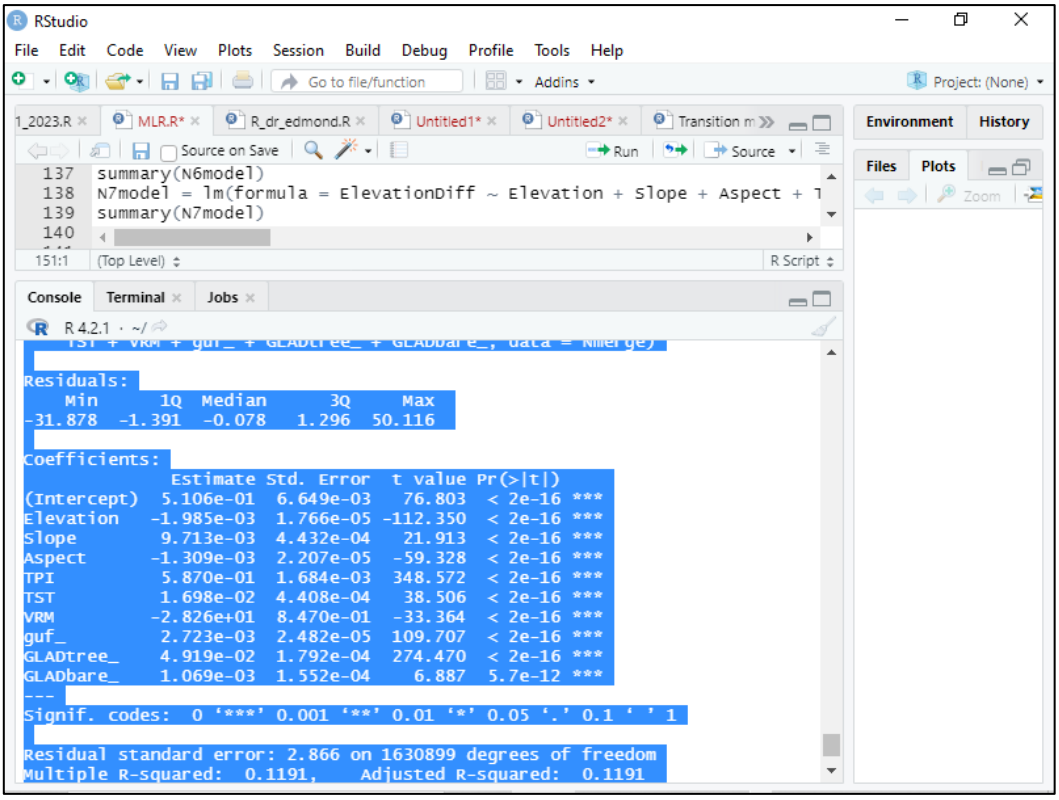


Figure 4.51: Execution of MLR using R in the Rstudio environment

4.11.6 DEM correction

Following the derivation of model coefficients, the regression models were adopted for correcting Copernicus and AW3D DEMs at the implementation sites shown in Figure 4.48. The corrected elevations were achieved by subtracting the predicted elevation error from the original elevations (i.e.., $\mathrm{DEM}_{\mathrm{Corrected}} = \mathrm{DEM}_{\mathrm{Original}} - \Delta h$).

4.12 Chapter Summary

The general and specific workflows in the methodology were addressed in this chapter. Section 4.1 summarised the overall methodology and the stages involved, section 4.2 discussed the characteristics of the various landscapes within the study area. In section 4.3, the datasets, terrain and land cover parameters were described. Sections 4.4, 4.5 and 4.7 discussed the data alignment techniques (i.e., mosaicking, datum harmonisation, co-registration and resampling), whereas the quality assessment of the DEMs is described in section 4.6. Sections 4.8 and 4.9 discussed the enhancement strategies while sections 4.10 and 4.11 discussed the performance evaluations and comparative assessments which were necessary for establishing the utility of the proposed DEM enhancement scheme and its proposal as a veritable strategy to the wider remote sensing community.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Introduction

Section 5.2 details the systematic evaluation of the qualities (and vertical accuracies) of the four global DEMs using both qualitative and quantitative measures, while section 5.3 addresses the results emanating from the comparative analysis of ten tree-based ensemble machine learning algorithms for DEM enhancement. Finally, the proposed enhancement framework is presented in section 5.4 while performance evaluations and comparative assessments are demonstrated in section 5.5.

5.2 DEM Quality Analysis

For a comprehensive assessment, several accuracy measures for both normal and non-normal data distributions are adopted, and the influence of terrain parameters on the variation of DEM error is also considered. Moreover, terrain and morphological attributes are also considered in the error analysis.

5.2.1 Analysis of the elevation distribution

The relationship between the global DEMs and LiDAR (after outlier filtering) is represented with the use of scatterplot matrices in Figure 5.1 - Figure 5.5 respectively. The lower panel, diagonal panel and upper panels display the scatter plots, histograms of the elevation distribution, and correlation values between the global DEMs and the 2 m LiDAR DEM elevations respectively. Generally, there is a close fit between the elevations, indicating strong positive correlations.

In the urban/industrial (Figure 5.1) and agricultural (Figure 5.2) landscapes, Copernicus DEM exhibits the strongest positive correlations of 0.98 and 1.00 respectively. Within the same environment, the other DEMs also have a very strong positive association with the lowest correlation of 0.96 occurring between ASTER GDEM and LiDAR. All DEMs show a perfect (1.00) or near-perfect (0.99) correlation with LiDAR in the mountain (Figure 5.3), peninsula (Figure 5.4) and grassland/shrubland (Figure 5.5) landscapes.

The elevation distribution in the built-up area and agricultural landscapes are normal or fairly normal, whereas the mountain landscape exhibits a bimodal shape. The first peak with elevations reaching 800 m indicates the high points on Table Mountain. The mountain slopes

downward into the adjoining Camps Bay, and this explains the second peak with elevations reaching 300 – 400 m.

Table 5.1 shows the summary of the elevation ranges. The elevation ranges from 0 to 1590 m, while the mean elevation ranges from 116 – 124 m with the lowest mean occurring in ASTER GDEM. Some of the highest points in Cape Town are in the eastern suburbs around the Lourensford dam, Table Mountain and Chapmans Peak.

Table 5.1: Summary of the elevation ranges

DEM	Min (m)	Max (m)	Mean (m)
LiDAR 2 m	0.00	1589.18	120.60
NASADEM 30 m	0.00	1590.51	121.46
ASTER 30 m	0.00	1578.51	116.04
Copernicus 30 m	0.00	1586.96	121.42
AW3D 30 m	0.00	1586.51	123.52

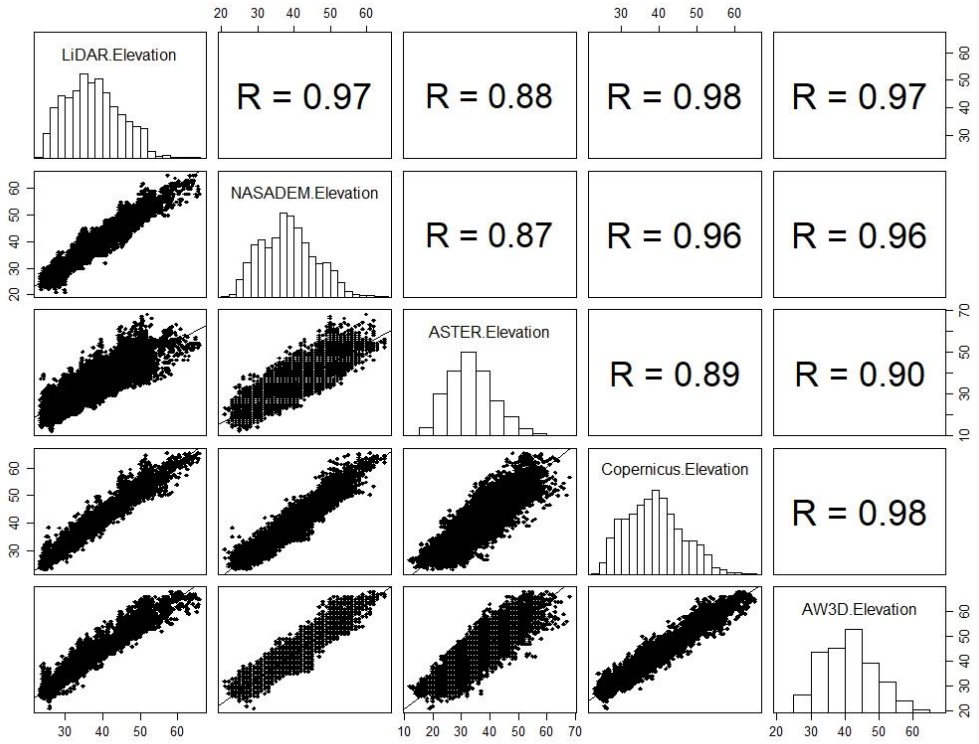


Figure 5.1: Graphical representation of the elevation distribution (diagonal panel), global DEM – LiDAR scatter plot (lower panel) and correlations (upper panel) in the Cape Town urban/industrial area

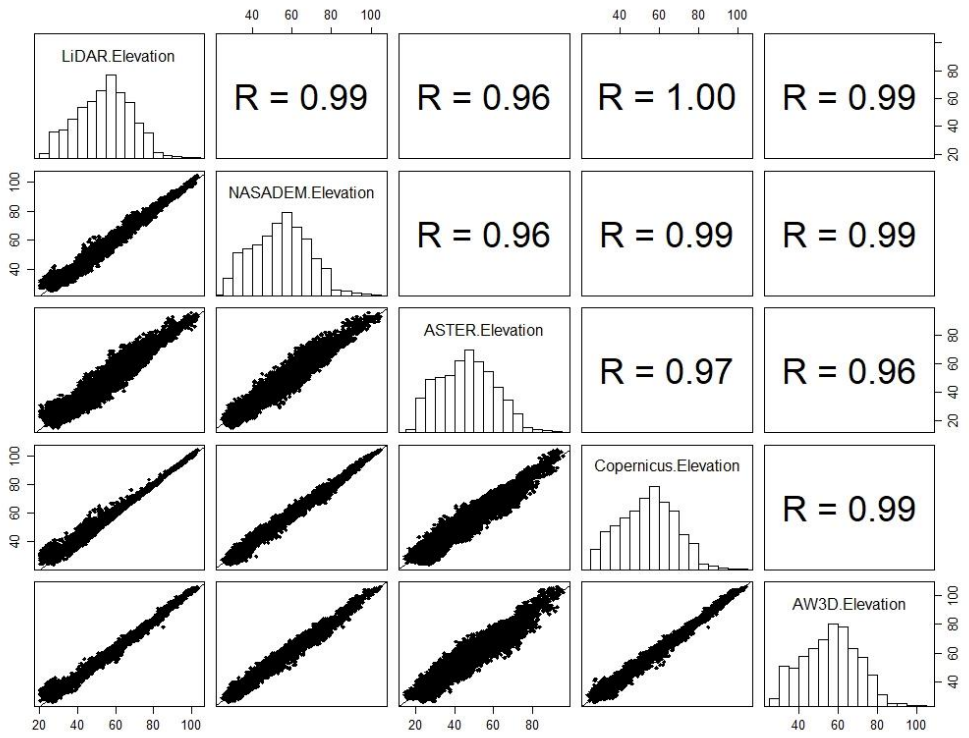


Figure 5.2: Graphical representation of the elevation distribution (diagonal panel), global DEM – LiDAR scatter plot (lower panel) and correlations (upper panel) in some agricultural lands of Cape Town

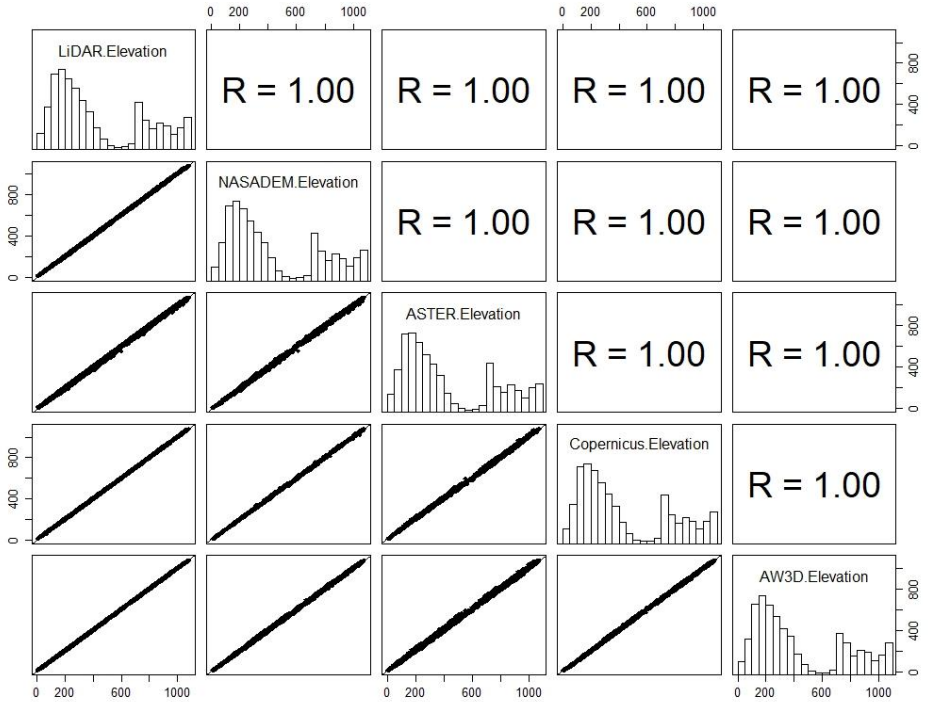


Figure 5.3: Graphical representation of the elevation distribution (diagonal panel), global DEM – LiDAR scatter plot (lower panel) and correlations (upper panel) in the Table Mountain and adjoining Camps Bay

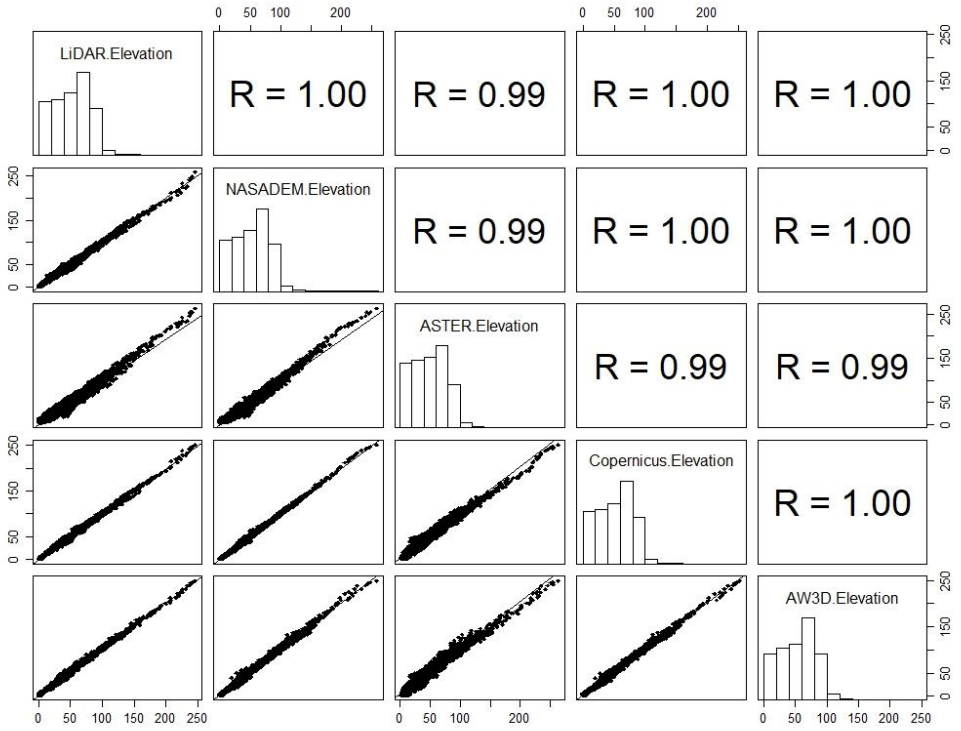


Figure 5.4: Graphical representation of the elevation distribution (diagonal panel), global DEM – LiDAR scatter plot (lower panel) and correlations (upper panel) in the Cape Peninsula

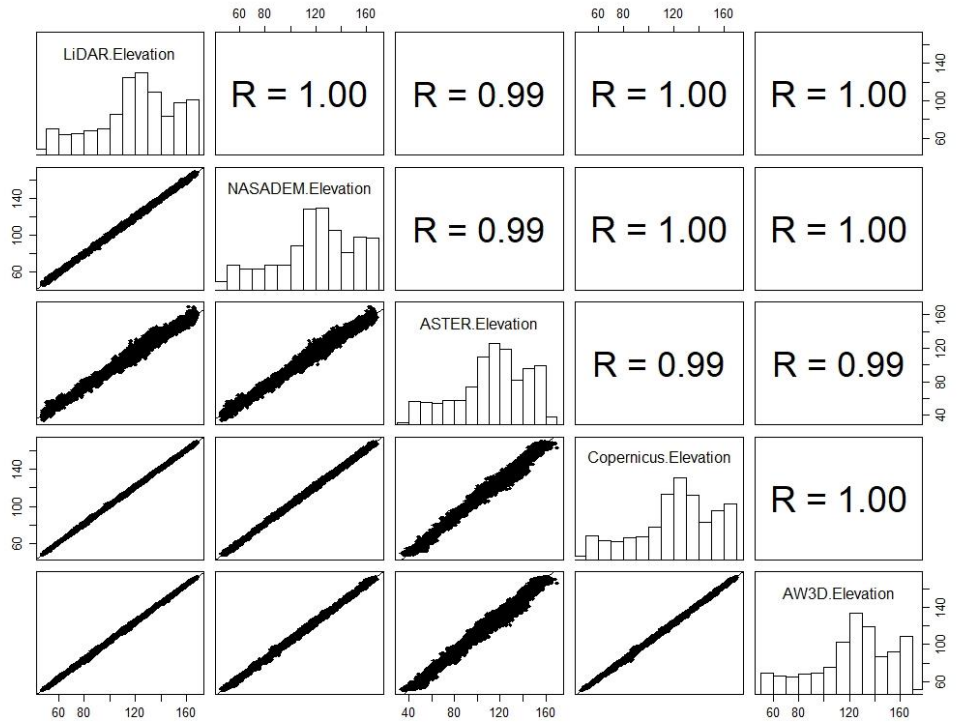


Figure 5.5: Graphical representation of the elevation distribution (diagonal panel), global DEM – LiDAR scatter plot (lower panel) and correlations (upper panel) in selected grasslands/shrublands/dense thickets

5.2.2 Elevation error distribution and analysis

Figures 5.6 – 5.15 present scatter plots of the elevation error before and after the outlier filtering rule (Three-sigma) was applied. Most errors in the urban/industrial area (Figure 5.6) are within the range of 23 – 60 m. Copernicus DEM exhibits the least error dispersion in the same urban/industrial area, unlike ASTER and AW3D. According to Olajubu et al. (2021), buildings are known to obscure the penetration of radar signals on the land surface. In the agricultural lands (Figure 5.8), the photogrammetric DEMs (ASTER and AW3D) have widely dispersed elevation errors, whereas there is a narrower error dispersion in the two SAR DEMs (NASADEM and Copernicus). It is known that DEM quality is affected by land cover and terrain steepness (Hoja et al., 2006). In Table Mountain and adjoining Camps Bay, this might explain the gradual escalation in elevation errors from between -50 to 50 m in the 0 – 400 m altitude region to as high as -100 m/+200 m at higher altitudes (Figure 5.10). The maximum positive elevation offsets in the mountain landscape occur in NASADEM and ASTER, whereas the maximum negative elevation offsets occur in Copernicus DEM. Summarily, the elevation error in the mountain landscape ranges from -185 to 230 m.

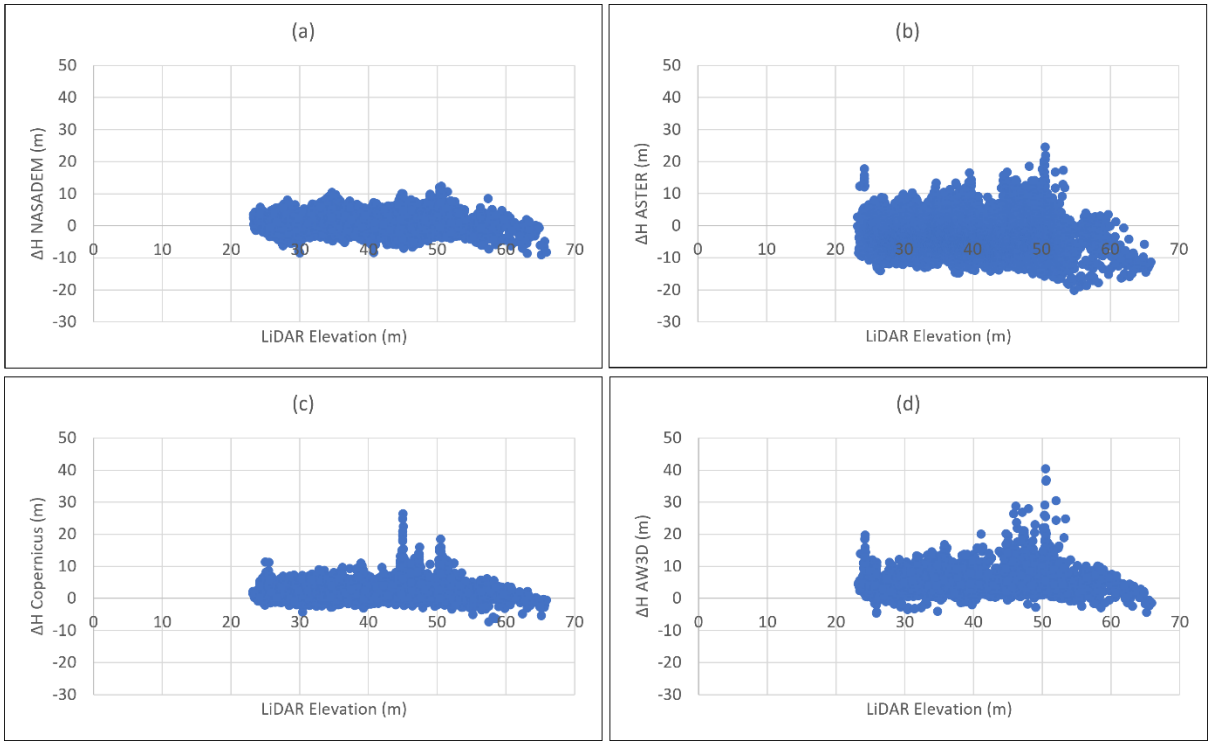


Figure 5.6: Graphical representation of the variations in elevation error, extracted from the raw data before Three-sigma outlier filtering – urban/industrial landscape

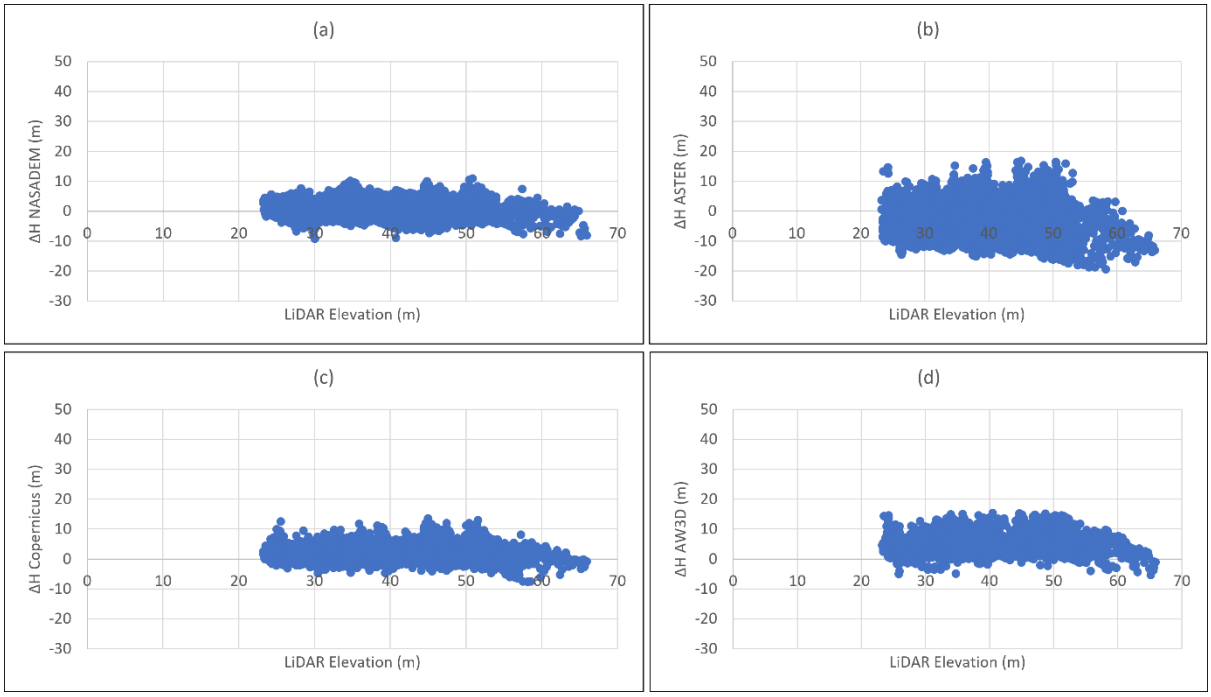


Figure 5.7: Graphical representation of the variations in elevation error, extracted from the data after Three-sigma outlier filtering – urban/industrial landscape

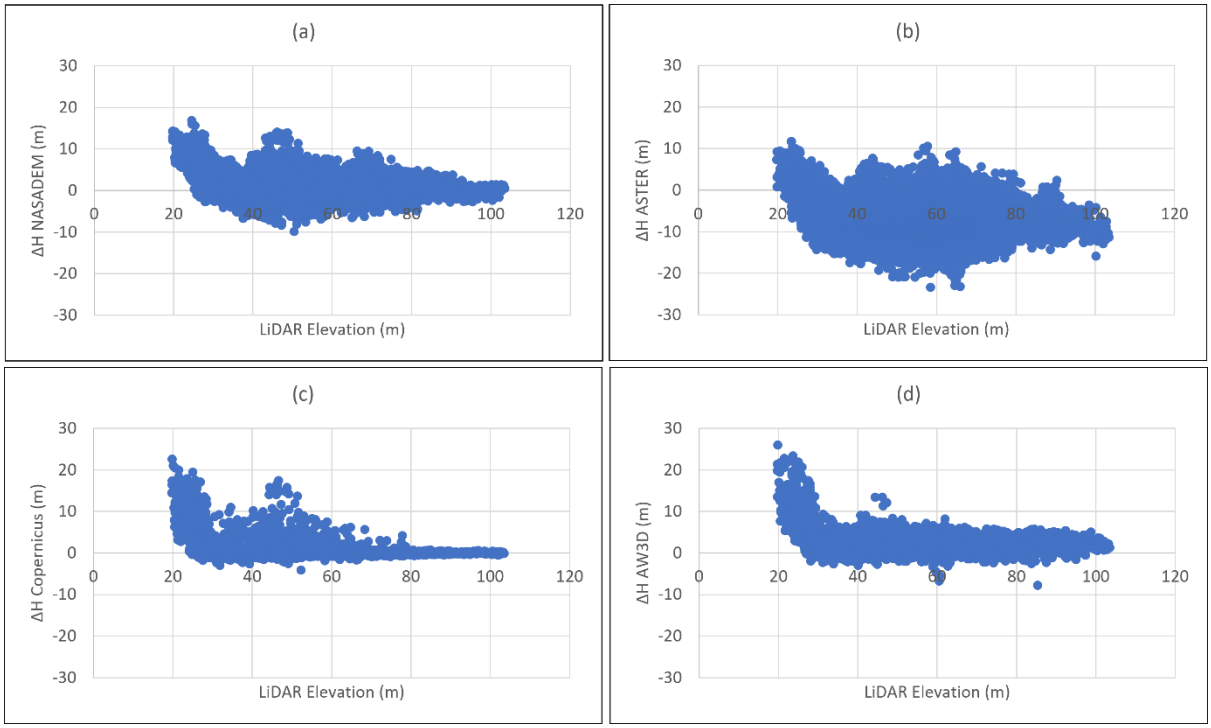


Figure 5.8: Graphical representation of the variations in elevation error, extracted from the raw data before Three-sigma outlier filtering – agricultural landscape

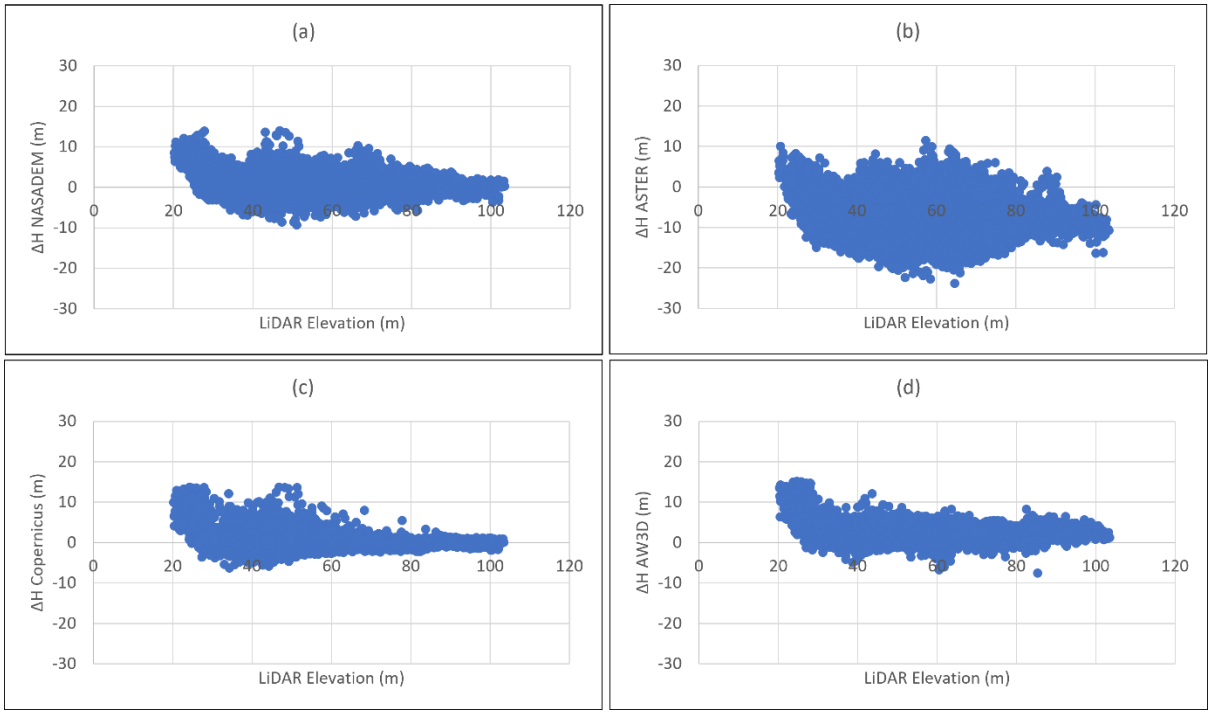


Figure 5.9: Graphical representation of the variations in elevation error, extracted from the data after Three-sigma outlier filtering – agricultural landscape

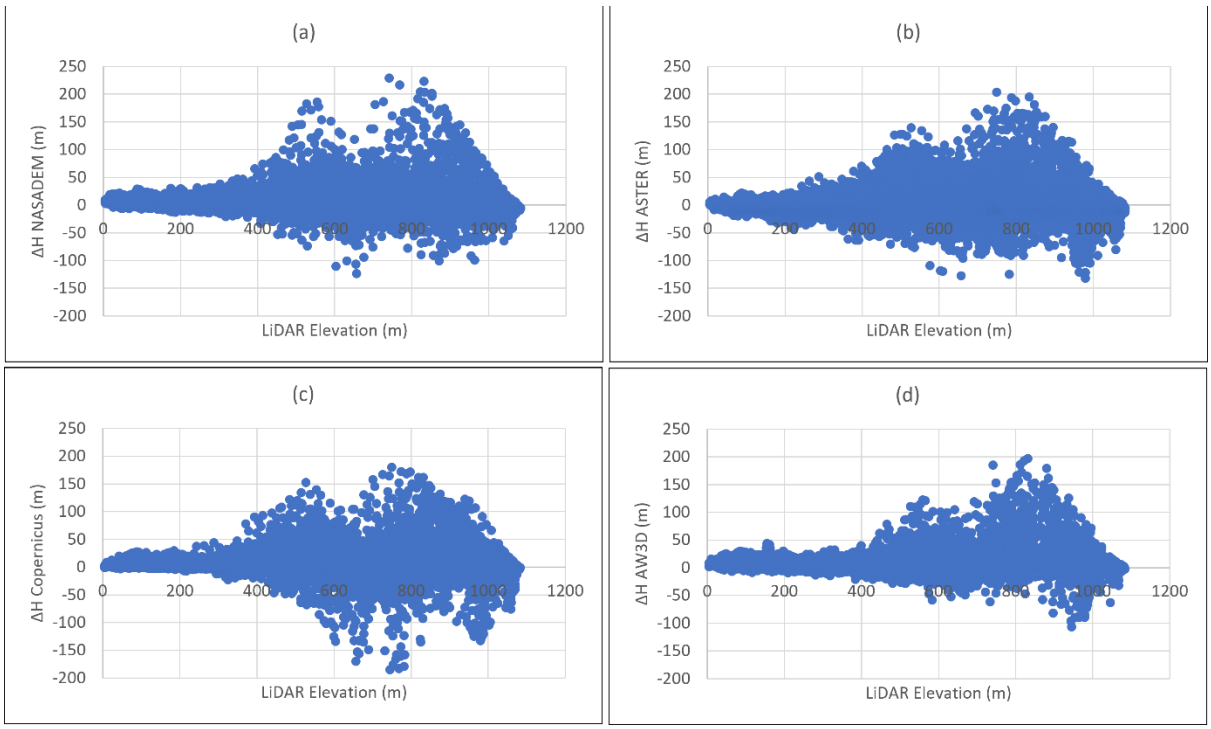


Figure 5.10: Graphical representation of the variations in elevation error, extracted from the raw data before Three-sigma outlier filtering – mountain landscape

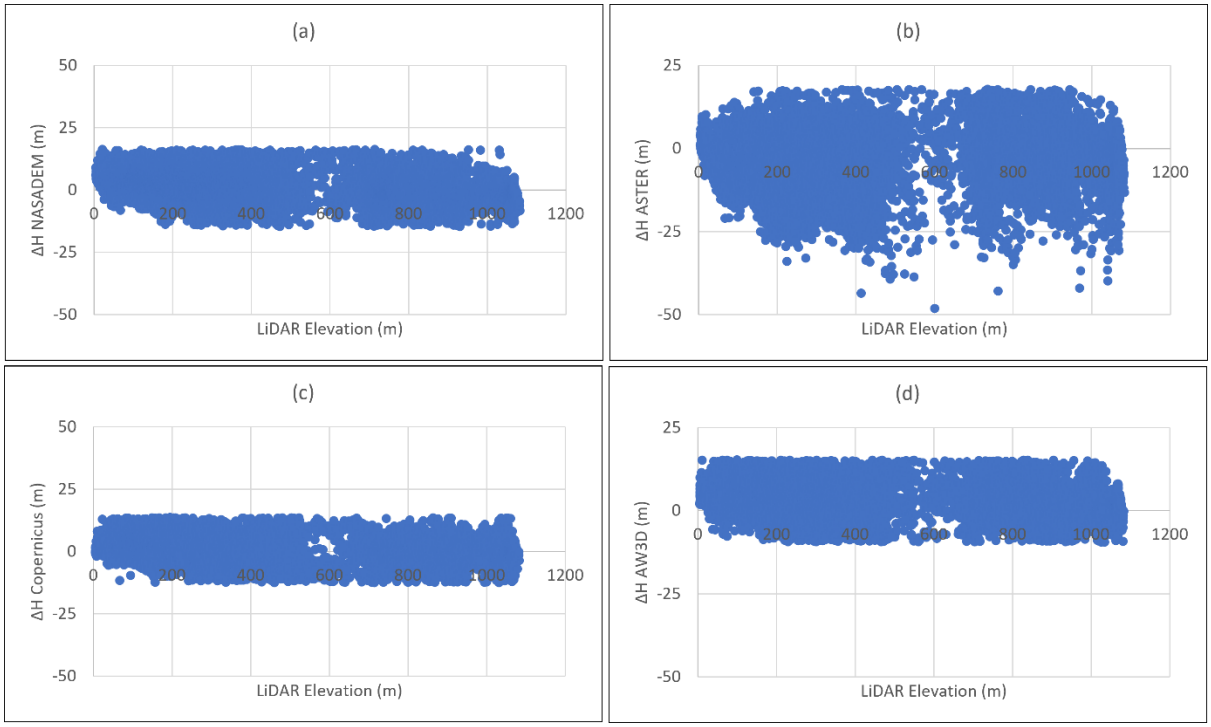


Figure 5.11: Graphical representation of the variations in elevation error, extracted from the data after Three-sigma outlier filtering – mountain landscape

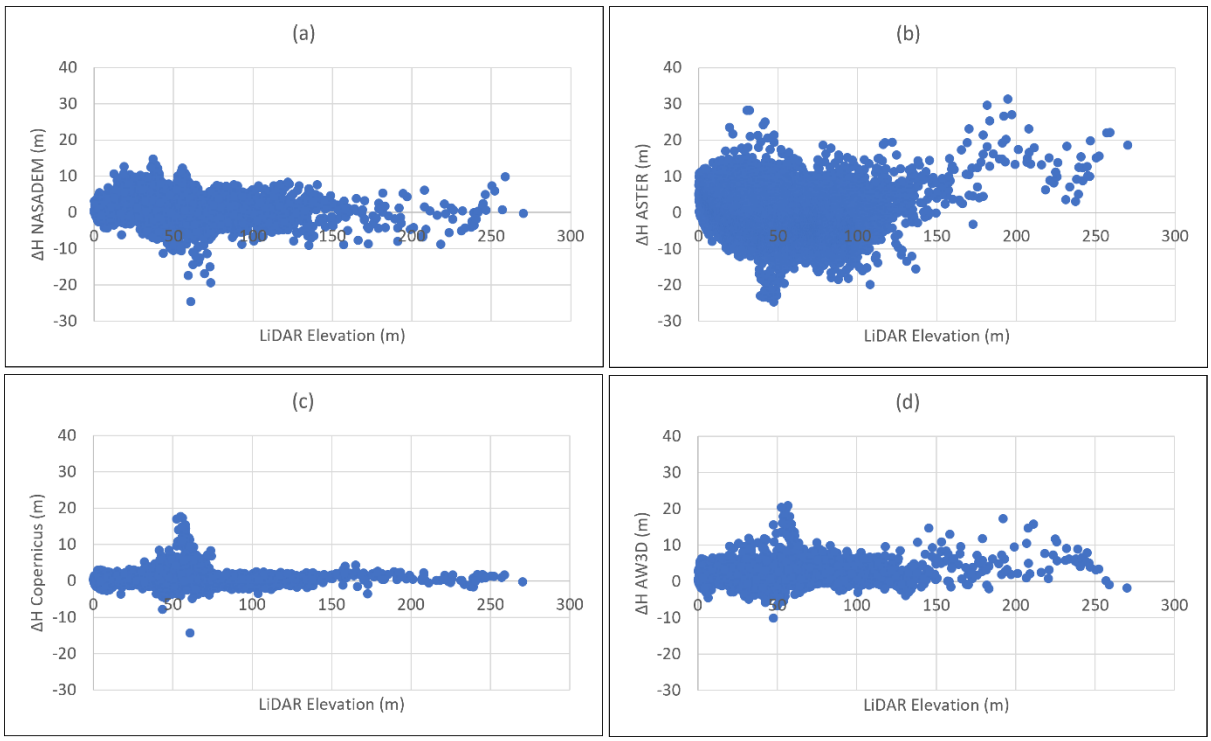


Figure 5.12: Graphical representation of the variations in elevation error, extracted from the raw data before Three-sigma outlier filtering – peninsula landscape

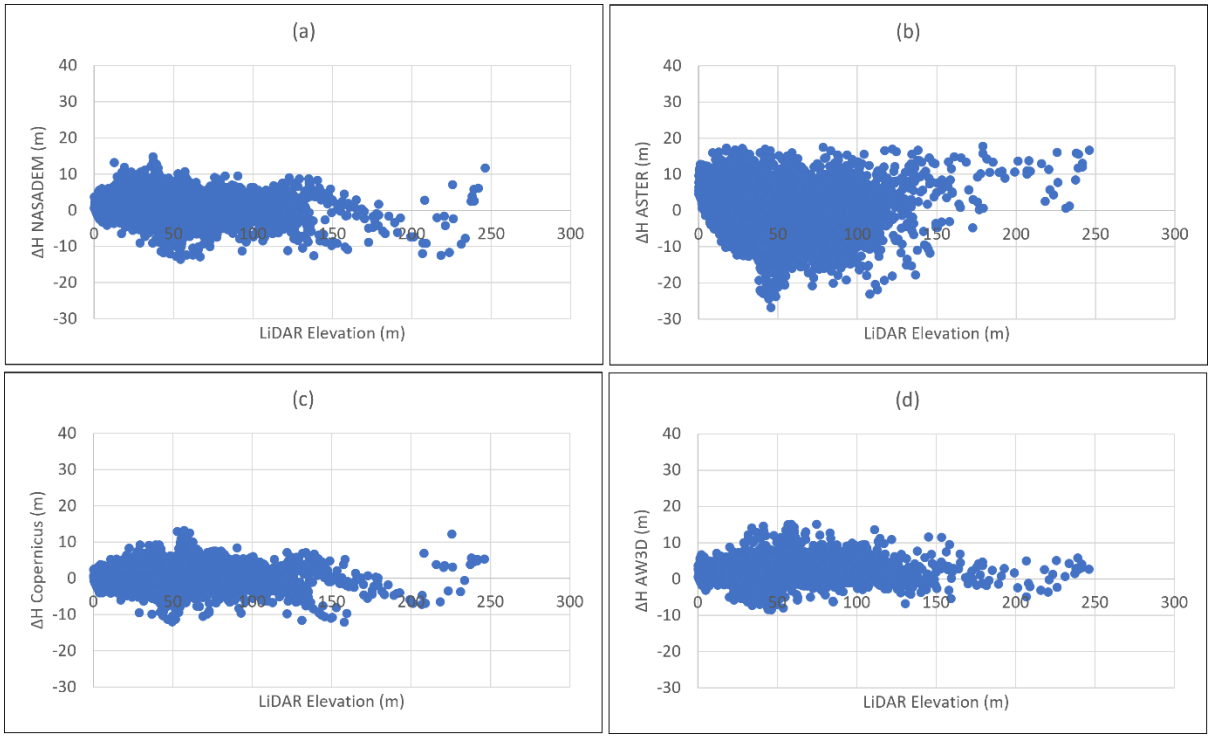


Figure 5.13: Graphical representation of the variations in elevation error, extracted from the data after Three-sigma outlier filtering – peninsula landscape

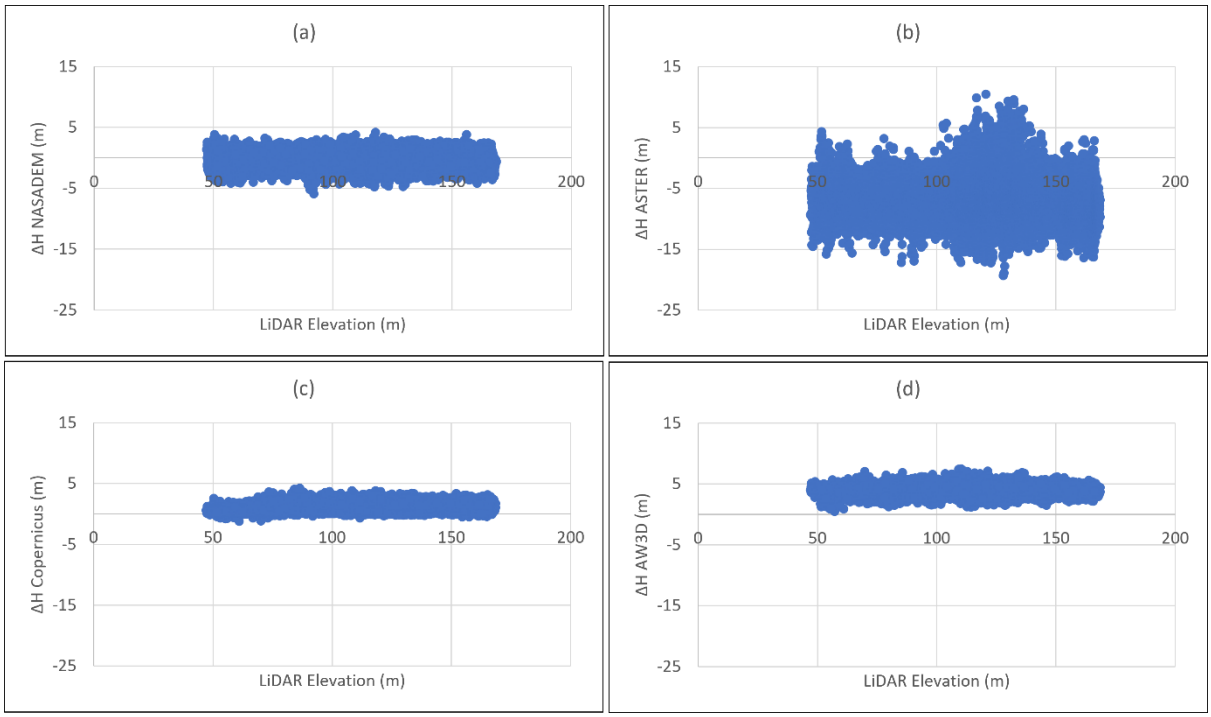


Figure 5.14: Graphical representation of the variations in elevation error, extracted from the raw data before Three-sigma outlier filtering – grassland/shrubland landscape

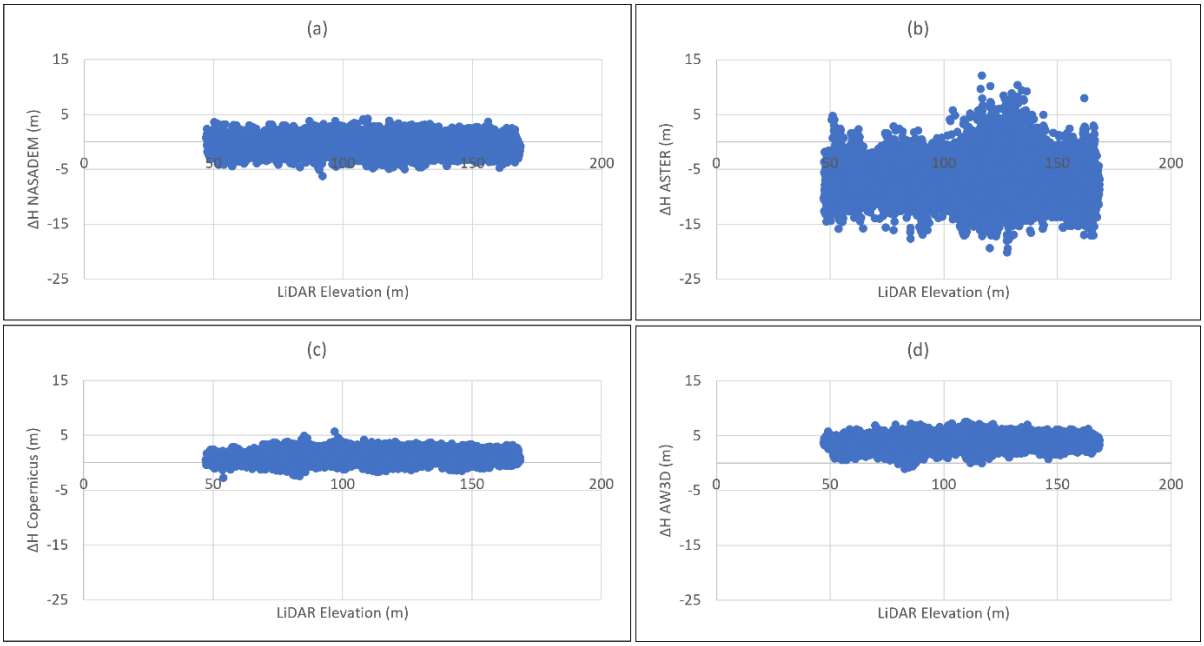


Figure 5.15: Graphical representation of the variations in elevation error, extracted from the data after Three-sigma outlier filtering – grassland/shrubland landscape

After outlier filtering, the histograms of the sampled errors (superimposed with curves for the normal distribution) are shown in Figure 5.16, and quantile-quantile (Q-Q) plots are shown in Figure 5.17. Positive elevation errors indicate areas where the DEMs overestimate the true surface elevation and negative elevation errors indicate areas where the true surface elevation is underestimated. There is a narrower spread of errors in Copernicus (-12.20 m to 14.70 m) and AW3D (-9.6 m to 15.6 m), and a greater range of errors in ASTER DEM (-53.9 m to 18.6 m).

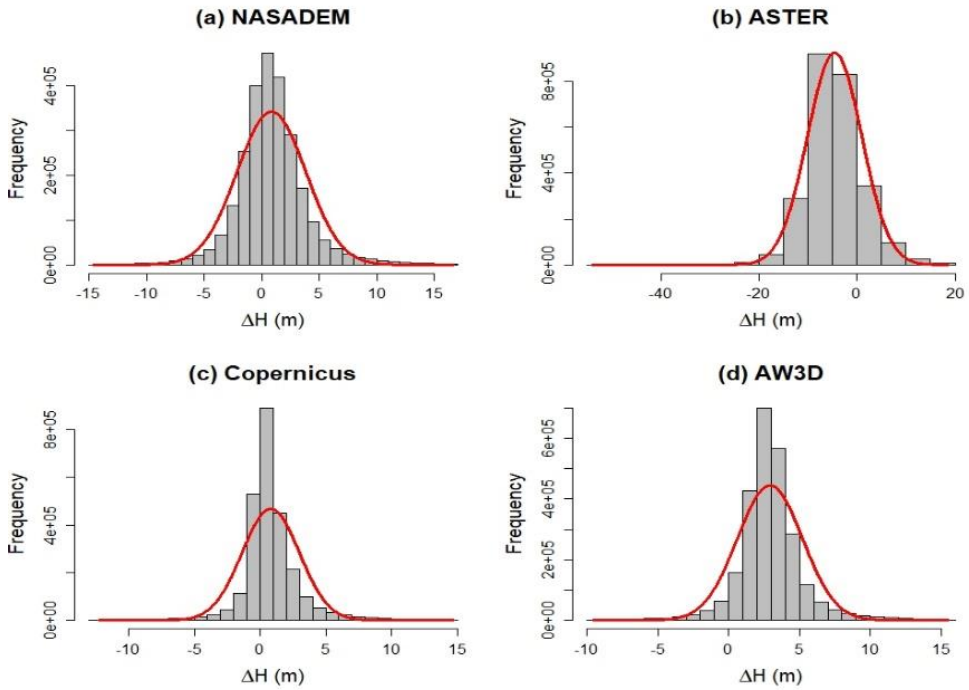


Figure 5.16: Histograms of all data points showing the distribution of ΔH in Cape Town, South Africa, (a) NASADEM (b) ASTER (c) Copernicus (d) AW3D, after outlier filtering

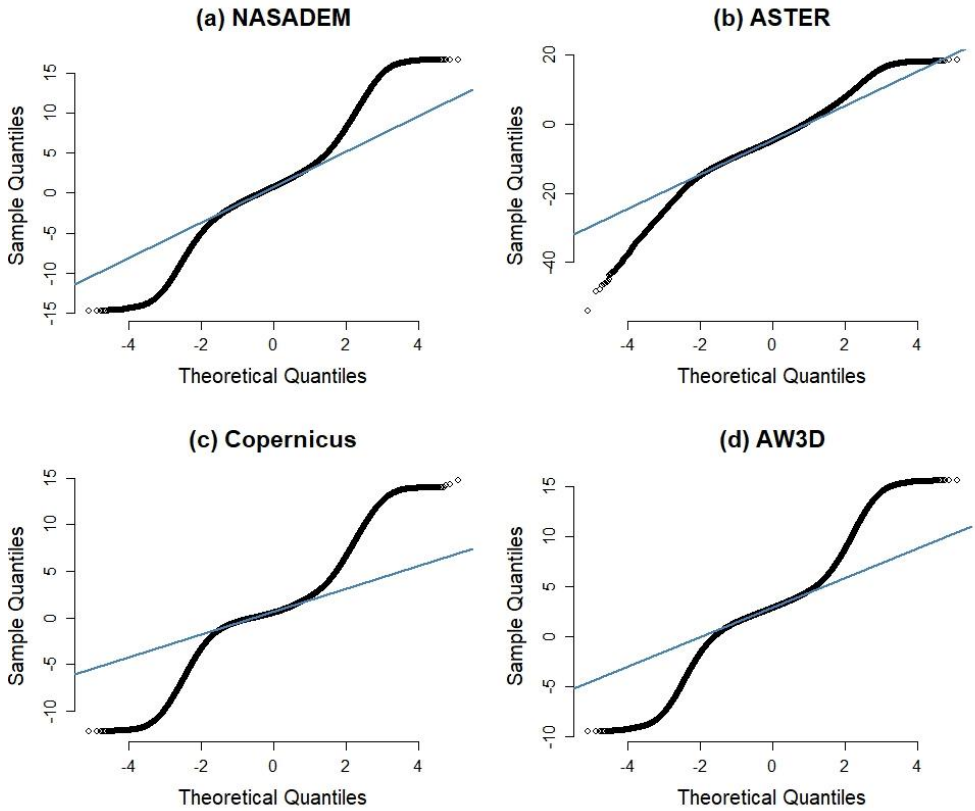


Figure 5.17: Normal Q-Q plot of all data points showing the distribution of ΔH in Cape Town, after outlier filtering

5.2.3 Visual and qualitative analysis

The visualisation of DEMs shown in Figure 5.18 and Figure 5.19 enhances the understanding of variations in the terrain characterisation from the multi-source DEMs. Airborne LiDAR provides the most detailed terrain conditioning and representation followed by Copernicus and AW3D in that order. In the agricultural landscape, the Diep River channel, its tributaries and their dendritic drainage pattern are visually perceptible in Copernicus and AW3D, whereas they are poorly distinguished in NASADEM and ASTER. Both NASADEM and ASTER display very coarse surface characterisation, unlike Copernicus and AW3D which exhibit finer hydrological details in the agricultural lands. Moderate elevations are more associated with residential areas while the industrial districts have higher elevations (e.g., Parow Industrial area). Also, the outlines and edges of building footprints are better discernible in Copernicus and AW3D but are diminished in NASADEM and ASTER.

Height error maps (or DEMs of difference) were generated and are shown in Figure 5.20. In the figure, negative offsets indicate points where the elevation is underestimated by the DEMs and vice versa. Urban and industrial areas, and the steep slopes of the Table Mountain and Cape Peninsula are some areas with considerably high offsets. This trend of increasing elevation errors in urban and industrial areas, and mountainous regions could be worsened by single or double bounce scatter during SAR data acquisition (Nwilo et al. 2022; Schlaffer et al. 2015; Tsyganskaya et al. 2016). Also, errors are induced in SAR measurements over steep mountain slopes due to shadow, layover and foreshortening (e.g., Wessel et al., 2018). However, measurement errors are less on the flat top of the Table Mountain which appears smoother. In a previous accuracy assessment of the global TanDEM-X DEM on the Table Mountain, Wessel et al. (2018) observed a similar scenario on the flat top of the mountain.

The very high positive and negative offsets along the steep edges of the Table mountain are worse in ASTER and NASADEM. In the urban areas and agricultural lands, the height error distribution is quite noisy in NASADEM and ASTER, whereas in Copernicus and AW3D, the higher negative offsets occur in the residential areas while the higher positive offsets occur in industrial areas. Similarly, in the agricultural lands, higher positive offsets are seen to occur along the Diep River channel, whereas higher negative offsets occur elsewhere along the floodplains.

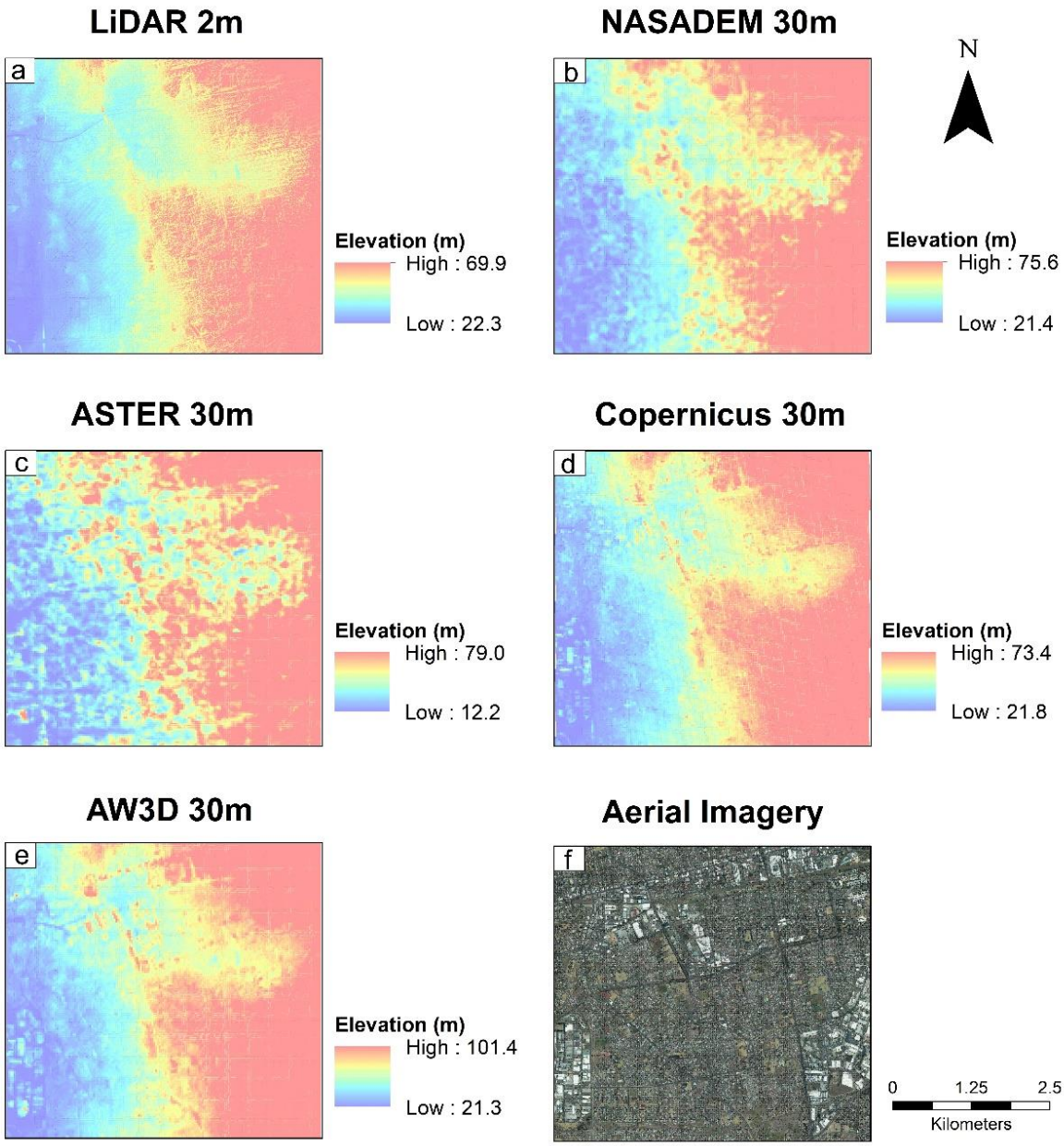


Figure 5.18: View of the LiDAR and global DEMs in the urban/industrial landscape, (a) LiDAR (b) NASADEM (c) ASTER (d) Copernicus (e) AW3D (f) January 2023 aerial imagery acquired from the City of Cape Town

5.2.4 Analysis of vertical accuracy

Table 5.2 shows the accuracy measures of the DEMs, and in Table 5.3 and Figure 5.21, the accuracy is categorised according to landscape. The positive mean errors of NASADEM, Copernicus and AW3D indicates their generally positive higher elevations which are systematically higher than the true ground surface. The converse is observed with ASTER which exhibited the highest negative mean error. Moreover, ASTER was the least precise in Cape Town (SD: 5.55 m) indicating a higher dispersion of the elevation errors, as evident in the histograms.

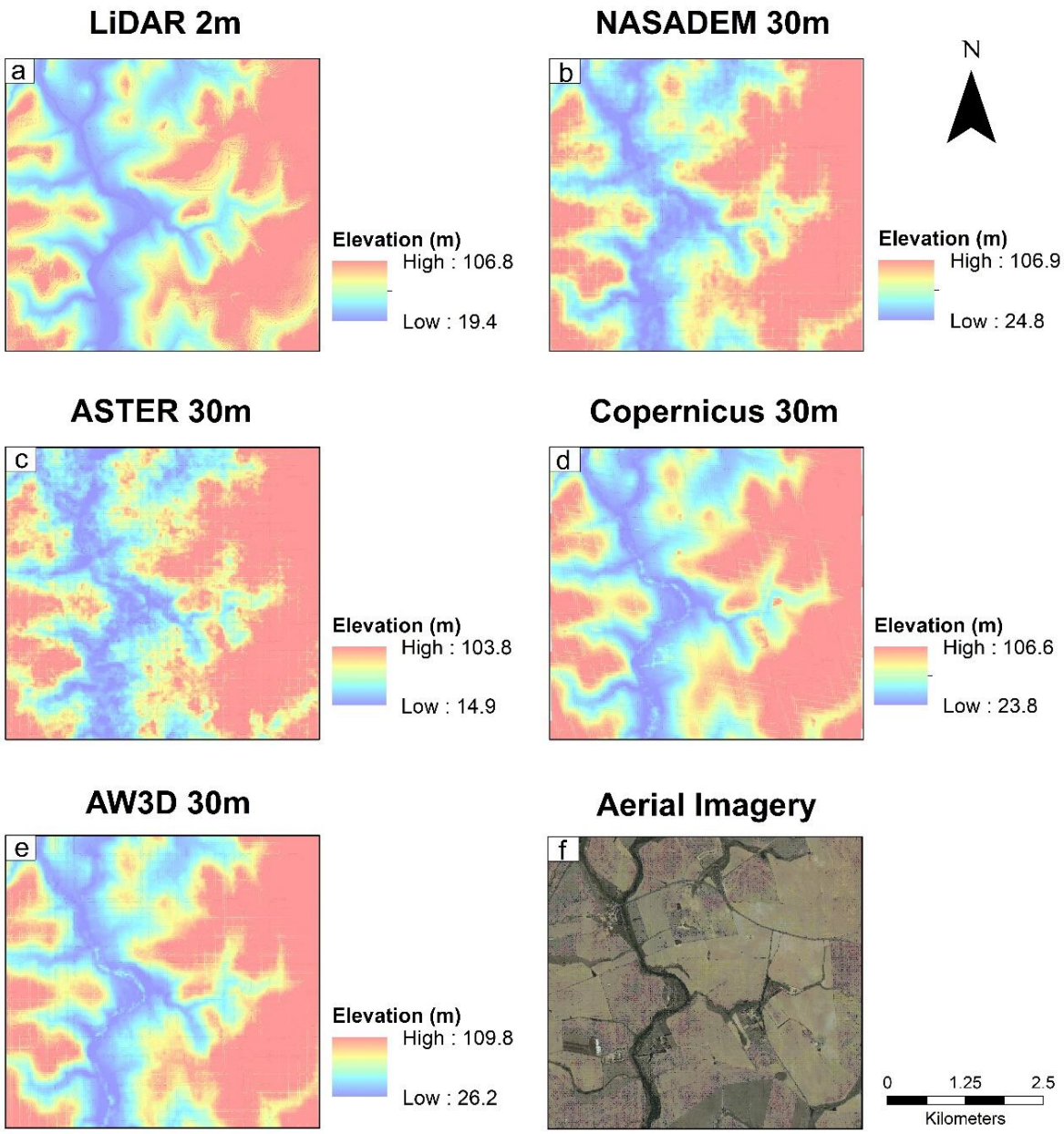


Figure 5.19: View of the LiDAR and global DEMs in the agricultural landscape, (a) LiDAR (b) NASADEM (c) ASTER (d) Copernicus (e) AW3D (f) January 2023 aerial imagery acquired from the City of Cape Town

AW3D and NASADEM had NMAD values of 1.47 m and 2.21 m respectively in Cape Town, occupying second and third place to Copernicus with an NMAD of 1.16 m. In Copernicus, the lowest RMSE of 1.36 m occurred in the Cape Peninsula, while the highest RMSE of 4.42 m occurred in the mountain landscape. Copernicus emerged as the DEM with the least vertical error, having the lowest mean error, SD, RMSE, MAE, NMAD, and linear error. Intermediately, NASADEM and AW3D occupied second and third place but with no clear distinction between them. For example, NASADEM had a lower mean error, RMSE and MAE

compared to AW3D. However, AW3D surpassed NASADEM with lower SD, NMAD and LE95.

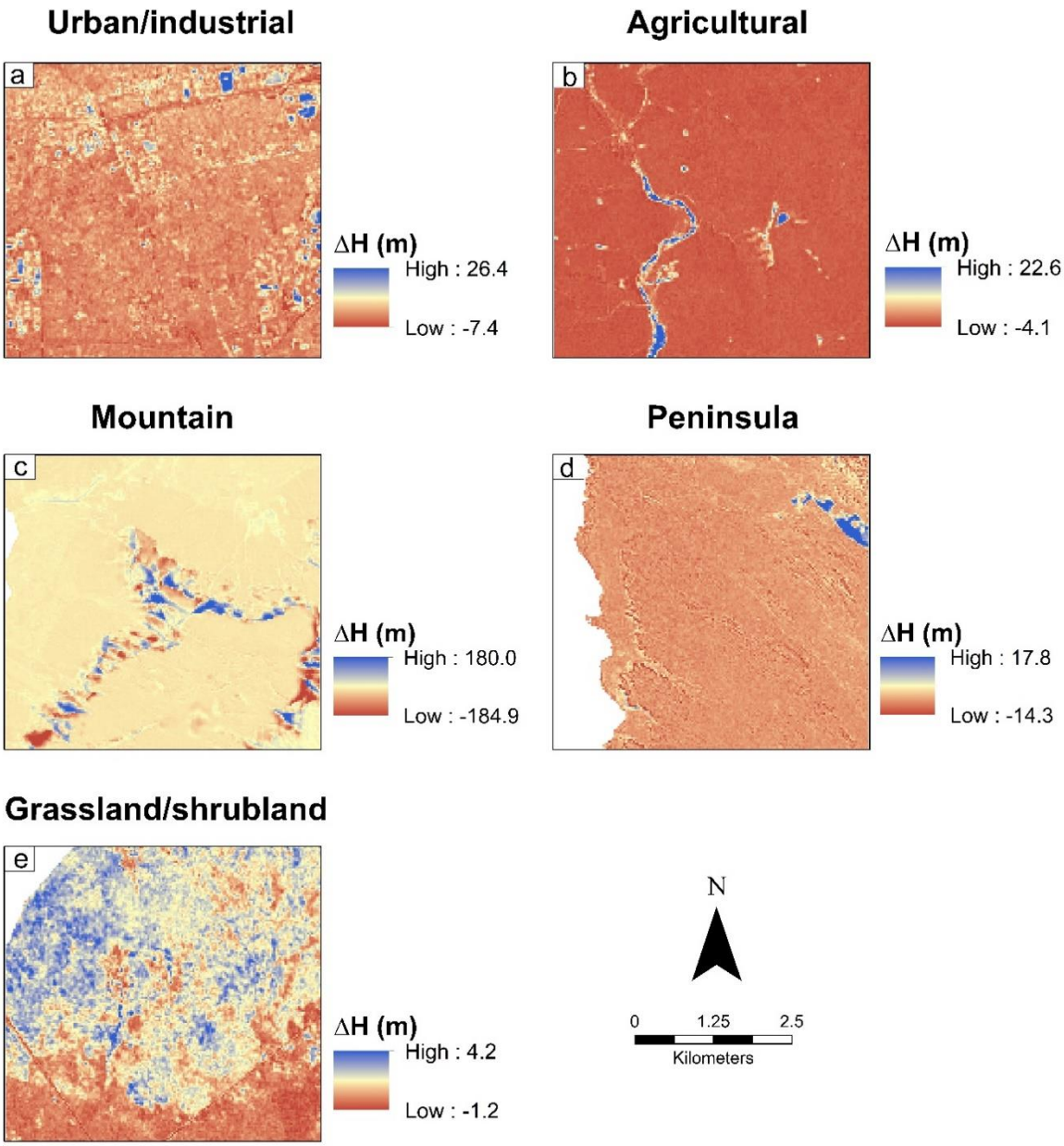


Figure 5.20: Height error maps of the Copernicus DEM at landscapes in Cape Town, (a) urban/industrial (b) agricultural (c) mountain (d) peninsula (e) grassland/shrubland

The vertical accuracy of ASTER GDEM was the poorest. Previous studies have shown the unsatisfactory performance of ASTER GDEM in several landscapes (e.g., Uuemaa et al., 2020). However, recent assessments have asserted the high vertical quality of Copernicus DEM (e.g., Purinton & Bookhagen, 2021; Li et al., 2022; del Valle et al., 2022). In summary, Copernicus DEM emerged with the least vertical error in all five landscapes.

Table 5.2: Vertical error for all sampled points. The lowest vertical errors are highlighted

DEM	**ME (m)**	**SD (m)**	**RMSE (m)**	**MAE (m)**	**MAD (m)**	**NMAD (m)**	**LE95 (m)**
NASADEM	0.86	3.01	3.13	2.20	1.49	2.21	5.89
ASTER	-4.56	5.55	7.18	5.94	3.32	4.93	10.87
Copernicus	**0.82**	**2.19**	**2.34**	**1.47**	**0.78**	**1.16**	**4.29**
AW3D	2.92	2.31	3.72	3.15	0.99	1.47	4.52

Table 5.3: Vertical error based on landscape. The lowest vertical errors are highlighted

Landscape	**DEM**	**ME (m)**	**SD (m)**	**RMSE (m)**	**MAE (m)**	**MAD (m)**	**NMAD (m)**	**LE95 (m)**
Urban/ industrial	NASADEM	1.35	1.85	2.29	1.85	1.13	1.67	3.63
	ASTER	-3.60	3.86	5.28	4.41	2.43	3.60	7.57
	Copernicus	**1.29**	**1.46**	**1.95**	**1.41**	**0.65**	**0.96**	**2.86**
	AW3D	4.39	1.89	4.78	4.40	0.83	1.23	3.71
Agricultural	NASADEM	0.51	2.12	2.18	1.59	1.18	1.76	4.15
	ASTER	-7.45	3.92	8.41	7.63	2.43	3.61	7.67
	Copernicus	**0.17**	**1.41**	**1.42**	**0.74**	**0.42**	**0.62**	**2.76**
	AW3D	2.37	1.53	2.82	2.42	0.78	1.16	3.00
Mountain	NASADEM	1.57	5.11	5.35	4.19	3.23	4.78	10.02
	ASTER	-4.01	7.65	8.63	6.67	4.47	6.63	14.99
	Copernicus	**1.13**	**4.27**	**4.42**	**3.40**	**2.56**	**3.79**	**8.37**
	AW3D	3.33	4.33	5.47	4.41	2.59	3.83	8.49
Peninsula	NASADEM	0.99	2.03	2.26	1.68	1.07	1.58	3.97
	ASTER	-2.29	4.85	5.36	4.40	3.01	4.46	9.51
	Copernicus	**0.17**	**1.35**	**1.36**	**0.80**	**0.46**	**0.68**	**2.64**
	AW3D	2.14	1.44	2.58	2.24	0.57	0.85	2.83
Grassland/ Shrubland	NASADEM	**-0.36**	1.15	**1.20**	**0.95**	0.76	1.12	2.25
	ASTER	-6.83	3.21	7.55	6.96	1.99	2.94	6.30
	Copernicus	1.36	**0.81**	1.58	1.40	0.57	0.84	**1.59**
	AW3D	4.09	0.82	4.17	4.09	**0.52**	**0.77**	1.61

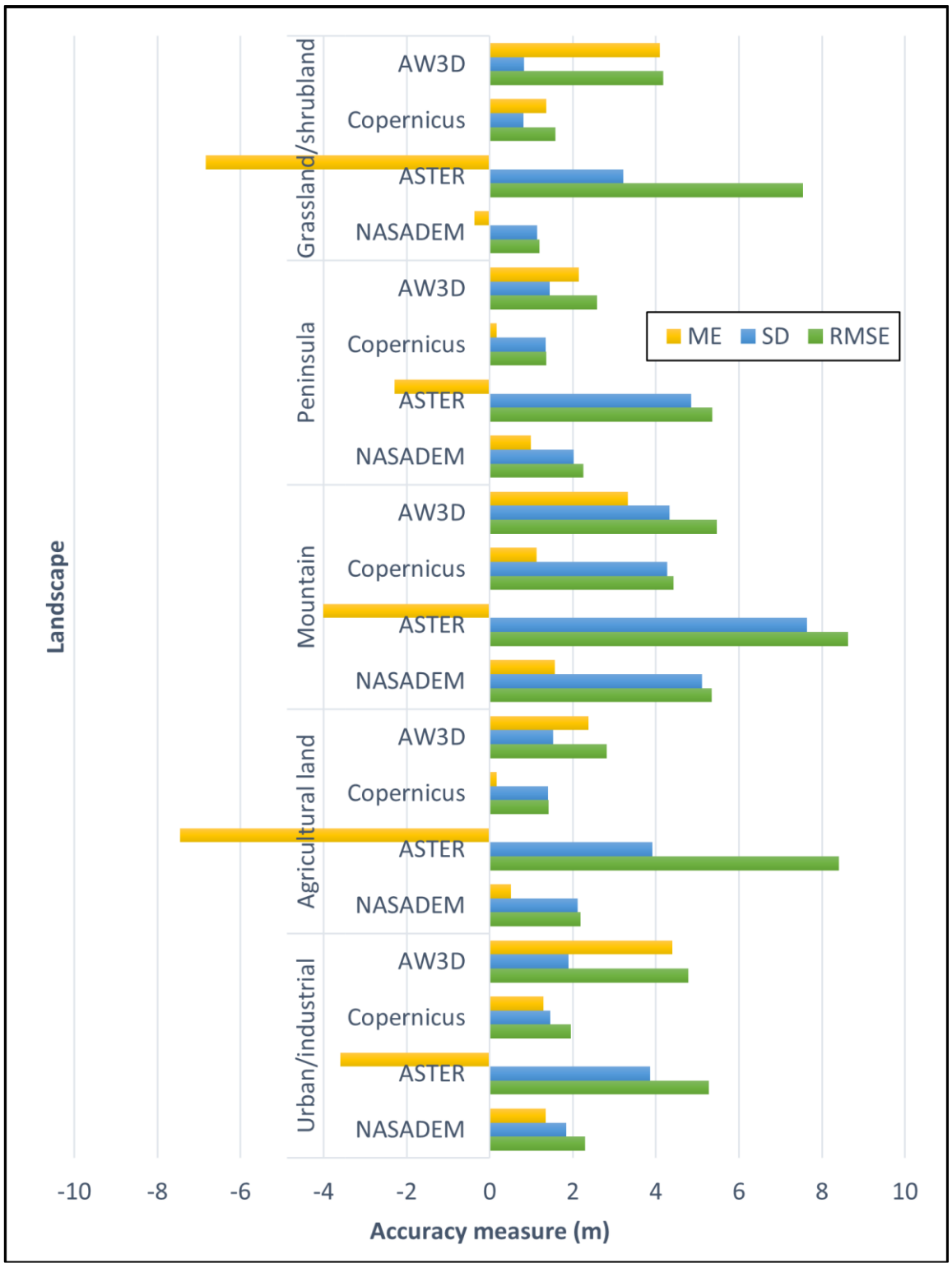


Figure 5.21: Visualisation of the mean error, SD and RMSE in different landscapes

Figure 5.22 compares the correlations between the elevation errors from the four DEMs. It shows that the elevation errors tend to have a low to moderate positive correlation in the mountains and urban areas.

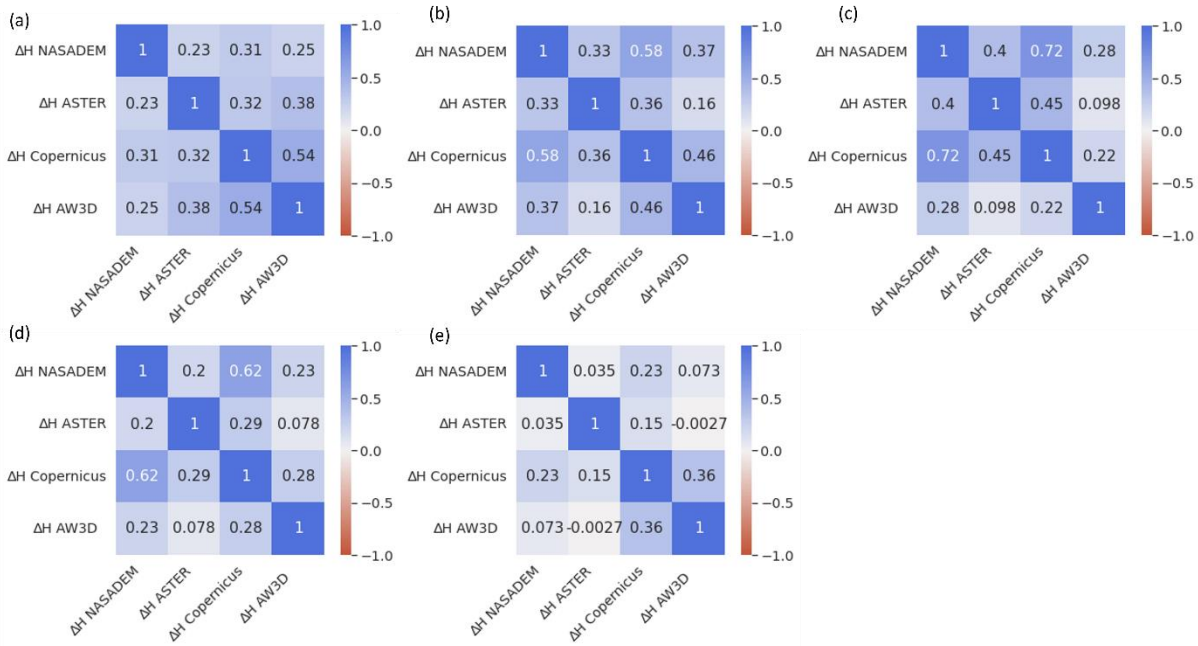


Figure 5.22: Correlations between the elevation errors of the four DEMs, after outlier filtering (a) Urban/industrial (b) Agricultural (c) Mountainous (d) Peninsula (e) Grassland/shrubland

5.2.5 Vertical accuracy by slope and aspect

Table 5.4 and Table 5.5 show the accuracy measures of the DEMs categorised by slope and aspect respectively. Generally, the DEMs perform better in the flat/gentle areas. However, as the slope progresses from flat/gentle to moderate, the SD, RMSE and MAE of Copernicus DEM almost increases by a factor of two. The general trend shows that the accuracies continue to decline as the slope gets steeper. Copernicus exhibits the least error while ASTER exhibits the highest error in most slope categories. Interestingly, the mean errors tend to be lower in steeper areas and higher in the gentle and moderate areas. Even the RMSE of Copernicus DEM in the steep, very steep, extremely and excessively steep slopes surpasses that of ASTER in the gentle slopes. Similarly, the ME, RMSE and MAE of NASADEM and AW3D in steep and very steep slopes perform better than ASTER in gentle slopes. These findings prove that the accuracy of ASTER DEM in Cape Town is not comparable to counterpart DEMs even within similar slope categories. Generally, the general trend shows a degradation in DEM accuracy in steeper slopes.

The lowest vertical errors in the DEMs are generally in the west direction, while some particularly higher errors are spotted in the northerly directions (north, northeast, and north-west). in the northerly direction (northeast, north and north-west). Copernicus emerges with the least vertical error having the lowest SD, RMSE and MAE in most directions, while ASTER

GDEM is the least accurate. Figure 5.23 and Figure 5.24 present a visualisation of the mean error, SD and RMSE based on slope and aspect respectively.

Table 5.4: Vertical accuracy categorised by the slope. The DEMs with the least vertical error are highlighted

Slope category	Description	DEM	ME (m)	SD (m)	RMSE (m)	MAE (m)
≤ 10°	Flat/ gentle	NASADEM	0.88	2.47	2.62	1.90
		ASTER	-4.59	4.93	6.74	5.68
		Copernicus	**0.83**	**1.64**	**1.83**	**1.17**
		AW3D	2.94	1.75	3.42	3.00
10° - 20°	Moderate	NASADEM	**0.70**	4.18	4.24	3.18
		ASTER	-4.60	6.91	8.30	6.71
		Copernicus	0.82	**3.16**	**3.27**	**2.40**
		AW3D	2.77	3.33	4.33	3.45
20° - 30°	Steep	NASADEM	0.76	5.51	5.56	4.36
		ASTER	-4.41	8.89	9.92	7.91
		Copernicus	**0.73**	**4.56**	**4.62**	**3.60**
		AW3D	2.81	4.73	5.50	4.39
30° - 40°	Very steep	NASADEM	0.79	6.37	6.42	5.17
		ASTER	-3.87	10.21	10.92	8.73
		Copernicus	**0.64**	**5.46**	**5.50**	**4.44**
		AW3D	2.99	5.57	6.32	5.13
40° - 50°	Extremely steep	NASADEM	0.66	6.91	6.94	5.67
		ASTER	-3.37	10.69	11.21	8.99
		Copernicus	**0.61**	**6.06**	**6.09**	**5.03**
		AW3D	3.09	6.11	6.85	5.65
> 50°	Excessively steep	NASADEM	**0.48**	7.42	7.43	6.16
		ASTER	-3.05	11.13	11.54	9.30
		Copernicus	0.88	6.52	**6.58**	**5.48**
		AW3D	3.36	**6.49**	7.30	6.08

Table 5.5 shows the accuracy measures of the DEMs categorised by aspect.

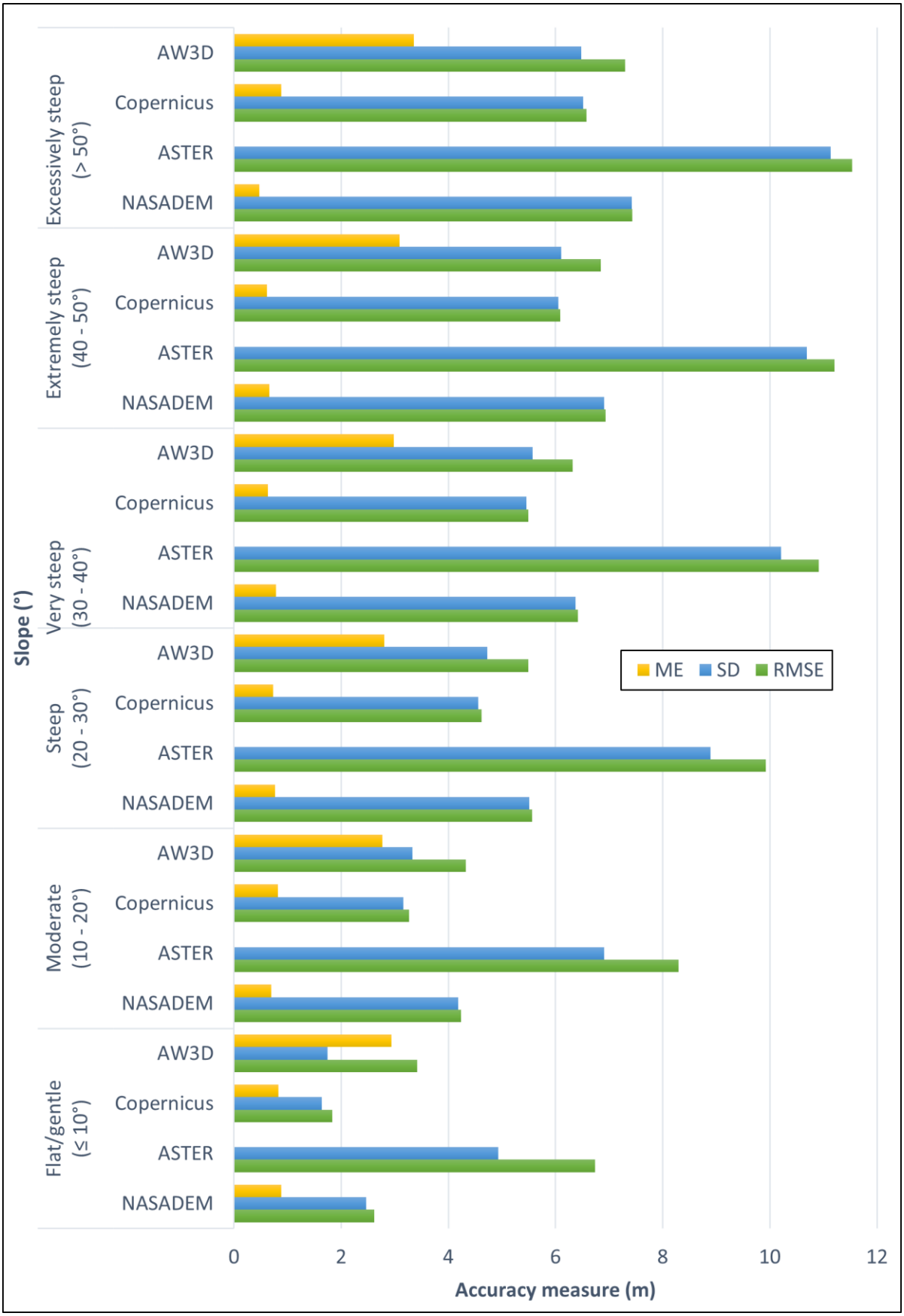


Figure 5.23: Visualisation of the mean error, SD and RMSE based on slope

Table 5.5: Vertical accuracy categorised by aspect. The DEMs with the least vertical error are highlighted

Aspect (°)	**DEM**	**ME (m)**	**SD (m)**	**RMSE (m)**	**MAE (m)**
North $(0-22.5°)$ $(337.5-360°)$	NASADEM	1.00	3.12	3.28	2.32
	ASTER	-5.64	5.57	7.93	6.58
	Copernicus	**0.78**	**2.33**	**2.46**	**1.53**
	AW3D	3.09	2.42	3.92	3.32
North-east $(22.5-67.5°)$	NASADEM	1.08	3.13	3.31	2.32
	ASTER	-5.60	5.78	8.05	6.61
	Copernicus	**0.71**	**2.27**	**2.38**	**1.50**
	AW3D	3.01	2.34	3.81	3.22
East $(67.5-112.5°)$	NASADEM	1.00	2.96	3.13	2.20
	ASTER	-5.19	5.74	7.74	6.36
	Copernicus	**0.81**	**2.12**	**2.27**	**1.43**
	AW3D	3.01	2.21	3.74	3.19
South-east $(112.5-157.5°)$	NASADEM	0.93	3.00	3.14	2.19
	ASTER	-4.96	5.38	7.31	6.10
	Copernicus	**0.90**	**2.15**	**2.33**	**1.47**
	AW3D	2.93	2.25	3.69	3.15
South $(157.5-202.5°)$	NASADEM	**0.83**	3.02	3.13	2.21
	ASTER	-3.96	5.28	6.60	5.53
	Copernicus	0.89	**2.23**	**2.41**	**1.53**
	AW3D	2.84	2.36	3.69	3.11
South-west $(202.5-247.5°)$	NASADEM	**0.67**	2.94	3.02	2.14
	ASTER	-3.15	5.44	6.28	5.22
	Copernicus	0.83	**2.17**	**2.32**	**1.48**
	AW3D	2.75	2.31	3.59	3.03
West $(247.5-292.5°)$	NASADEM	**0.68**	2.84	2.92	2.07
	ASTER	-3.85	5.30	6.55	5.46
	Copernicus	0.81	**2.03**	**2.19**	**1.37**
	AW3D	2.85	2.20	3.60	3.06
North-west $(292.5-337.5°)$	NASADEM	0.83	3.08	3.19	2.25
	ASTER	-5.17	5.41	7.48	6.26
	Copernicus	**0.82**	**2.27**	**2.41**	**1.49**
	AW3D	3.00	2.36	3.82	3.24

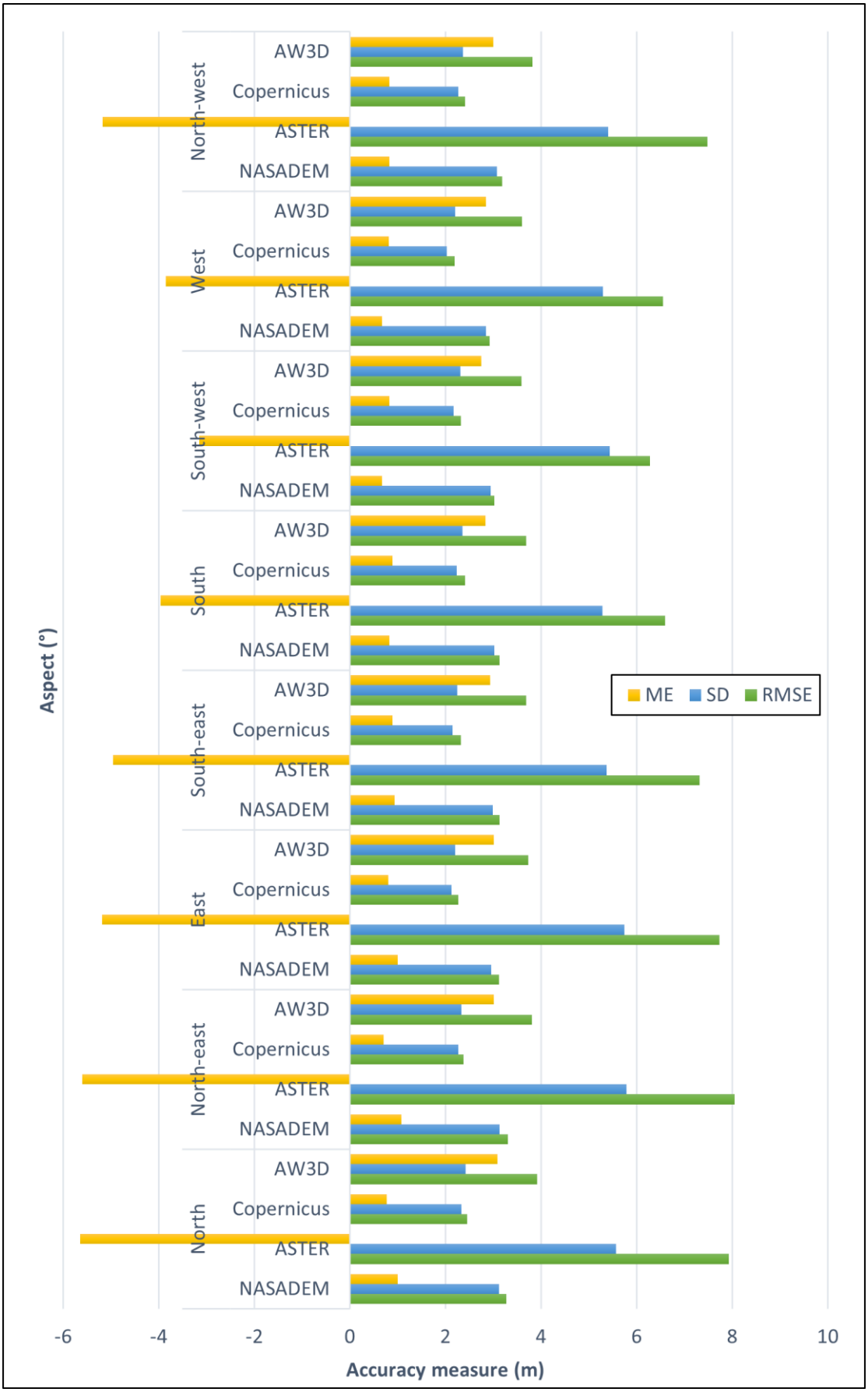


Figure 5.24: Visualisation of the mean error, SD and RMSE based on aspect

5.2.6 Spatial pattern of detected errors

Figure 5.25 presents the Global Moran's I summary of the spatial pattern of the elevation errors. Given the z-scores ranging from 2916.341 (Copernicus) to 5045.906 (ASTER), there is less than 1% likelihood that the clustered pattern of elevation errors could be the result of random chance.

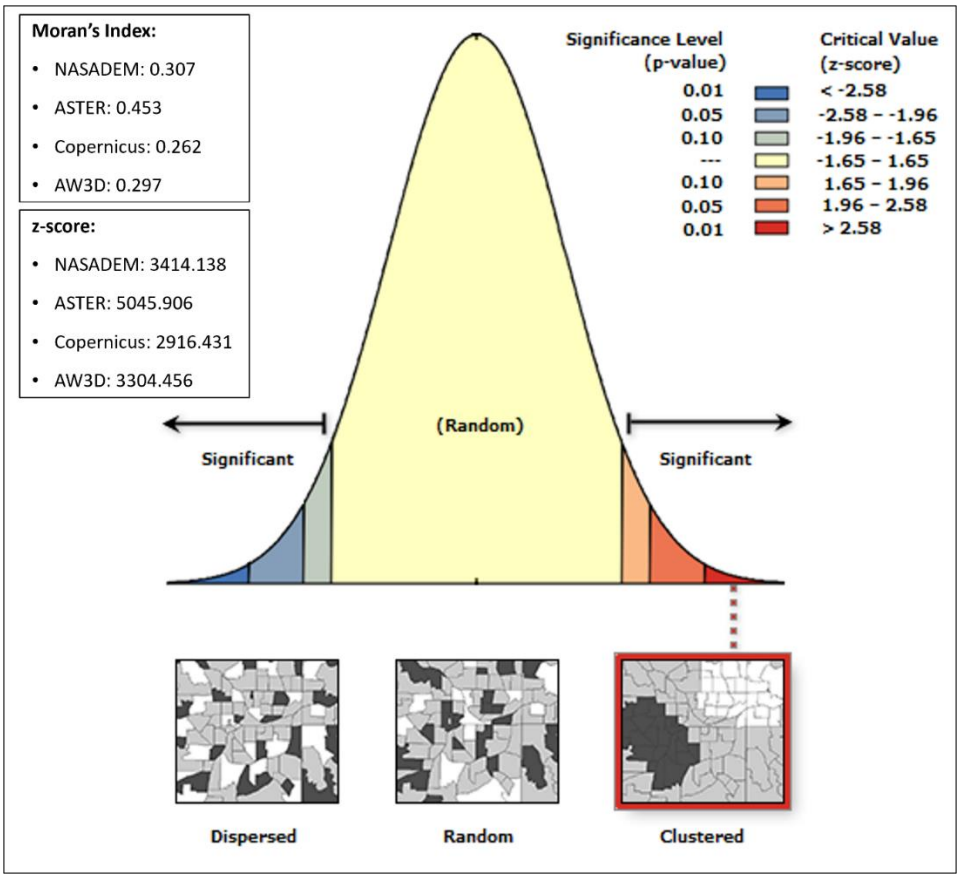


Figure 5.25: Global Moran's I summary

5.2.7 Comparison of geo-morphometric characteristics

There is a general consensus in the existing literature that the accuracy and resolution of DEMs influence their suitability for hydrological modelling (e.g., Tavares da Costa et al., 2019; Dávila-Hernández et al., 2022). More so, the resolution of DEMs influences the hydrological and hydraulic assessments of watersheds, with impacts on decision-making for watershed management and development (Tesema, 2021). The sub-basin shown in Figure 4.13 was selected for geo-morphometric analysis of the three DEMs with the least vertical errors (Copernicus, AW3D and NASADEM), with the exclusion of ASTER GDEM. Similarly, (Purinton & Bookhagen, 2017) excluded ASTER from their terrain analysis because of its elevation noise and vertical uncertainty. The LiDAR DEM was resampled from 2 m to 30 m to generate drainage networks which are comparable to the satellite DEMs, and for practical

purposes, we regard it as the reference for this evaluation. Table 5.6 presents the analysis of geo-morphometric parameters for the selected sub-basin. The analysis shows the highest consonance between the geo-morphometric parameters of Copernicus DEM and LiDAR.

Table 5.6: Analysis of geo-morphometric parameters for the selected sub-basin in Cape Town

	LiDAR 30 m (resampled)	**NASADEM 30 m**	**Copernicus 30 m**	**AW3D 30 m**
Drainage network parameters				
Total no. of stream order	151353.0	195754.0	159253.0	171364.0
Total length of streams (m)	20373996.6	17679004.3	19790459.6	17842469.7
Geometry parameters				
Total Basin Area(km^2)	673.3	670.4	670.4	670.4
Basin Length (km)	40.5	39.1	40.5	39.3
Main Channel Length (km)	79.8	72.0	74.2	70.3
Mean Basin Width	16.6	17.1	16.6	17.1
Drainage Texture	728.1	941.7	766.1	824.4
Drainage texture analysis				
Drainage Density (km/km^2)	30.3	26.4	29.5	26.6
Stream Frequency ($number/km^2$)	224.8	292.0	237.6	255.6
Drainage Intensity	7.4	11.1	8.0	9.6
Average Length of Overland Flow (km)	0.0	0.0	0.0	0.0
Basin relief				
Maximum Height of Basin(m)	456.0	460.0	457.0	461.0
Total Basin Relief (H) m	457.0	460.0	457.0	461.0

The highest stream frequency, drainage density and drainage intensity are from the LiDAR dataset. With a drainage density of 29.521 km/km^2, Copernicus DEM is closest to the LiDAR estimate while NASADEM is furthest with a drainage density of 26.37 km/km^2. This corresponds with the previous assessment in which Copernicus emerged with the least vertical error. As the coarseness of a DEM increases, its vertical accuracy decreases, and the vertical difference between adjacent pixels increases. This greatly increases the flow distance between adjacent pixels. Expectedly, coarser grid DEMs are less applicable for hydrological modelling in urban areas where the areas prone to flooding must be delineated more precisely. In the comparison, Copernicus DEM was comparable to the resampled LiDAR DEM in the average length of overland flow. This indicates it is a better option for wide-area hydrological modelling

in regions with mixed land uses. The differences in the drainage network could be influenced by the DEM or other factors such as the drainage network algorithm (Polidori & El Hage, 2020).

This analysis provides a perspective on the hydrological conditioning of the global DEMs, including their suitability for hydrological operations (e.g., flood modelling and watershed delineation). According to Muthusamy et al. (2021), coarser DEMs could reduce river channel conveyance. Moreover, topography is a crucial element in the hydrological response of a drainage basin (Tavares da Costa et al., 2019), and the delineation of watersheds is significantly affected by the DEM resolution (Al-Khafaji et al., 2020). Figure 5.26 shows the drainage network in the 6th – 9th stream order range for the selected sub-basin.

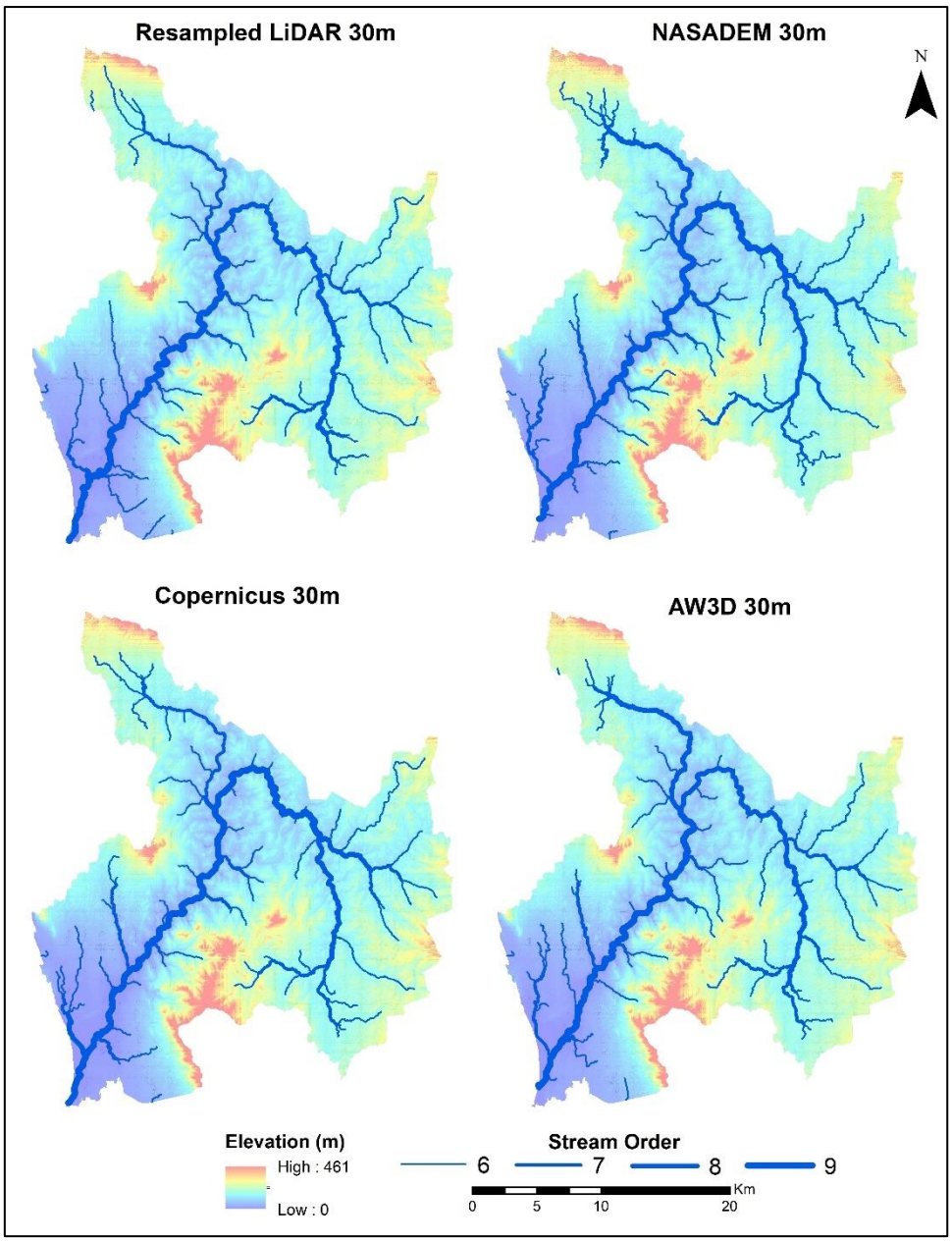


Figure 5.26: Drainage network in the 6th – 9th stream order range for the selected sub-basin in Cape Town

5.2.8 Discussion of DEM quality assessment results

The analysis has determined that Copernicus GLO-30 has the least vertical error in the urban/industrial, agricultural, mountain, peninsula and grassland/shrubland landscapes. Generally, the elevation errors from the individual DEMs have a low to moderate positive correlation in urban and mountainous areas. Judging by established indices, Copernicus DEM outperformed all other DEMs in all the assessed landscapes. ASTER GDEM's performance is below expectations, and shows considerable negative bias with a mean error of -4 m. This analysis has provided an extensive validation of NASADEM, ASTER, Copernicus and AW3D DEMs to guide practitioners on their respective limitations for numerous applications. The

multi-level analysis goes beyond the reliance on global uncertainty measures by analysing localised differences through terrain visualisation.

Several factors favour the selection of Copernicus and AW3D for the next stage of enhancement:

1. Copernicus had the least vertical error among the compared DEMs, and AW3D was a close runner-up as it surpasses NASADEM when robust accuracy measures such as the NMAD are considered.
2. Among the interferometric DEMs assessed (NASADEM and Copernicus), Copernicus emerged with the least vertical error. Similarly, among the photogrammetric DEMs (ASTER and AW3D), AW3D emerged with the least vertical error. This research prioritises the inclusion of multi-sensor and multi-source DEMs emanating from different production technologies for the next stage of enhancement.
3. Among the 30 m DEMs, the terrain representation by Copernicus and AW3D had the finest resolution, based on the qualitative (visual) analysis. Also, the outlines and edges of building footprints were better discernible in Copernicus and AW3D.
4. The data acquisition periods for Copernicus and AW3D are most recent for the study area, as shown in Table 4.1.

5.3 Evaluation of Tree-based Ensembles for DEM enhancement

5.3.1 Descriptive analysis of terrain parameters

Figure 5.27 and Figure 5.28 show the derived terrain parameters and height error map for the selected area. The streams weaving through the agricultural lands in the basin exhibit a dendritic drainage pattern, and the slope ranges from gentle to moderately steep, and steep along the south-eastern slopes (Figure 5.28a). The Mosselbank River located in the south-east empties into the main drainage channel of the Diep River which flows downwards in the south-west direction. The areas with steep slopes tend to exhibit higher TPI values where central points in certain neighbourhood are located higher than their average surroundings. Table 5.7 shows descriptive statistics of the predictor variables used for training/testing. Figure 5.29 presents histograms showing the elevation error (ΔH) distribution at the selected sites. Most of the elevation errors are in the range of -5 to +5 m.

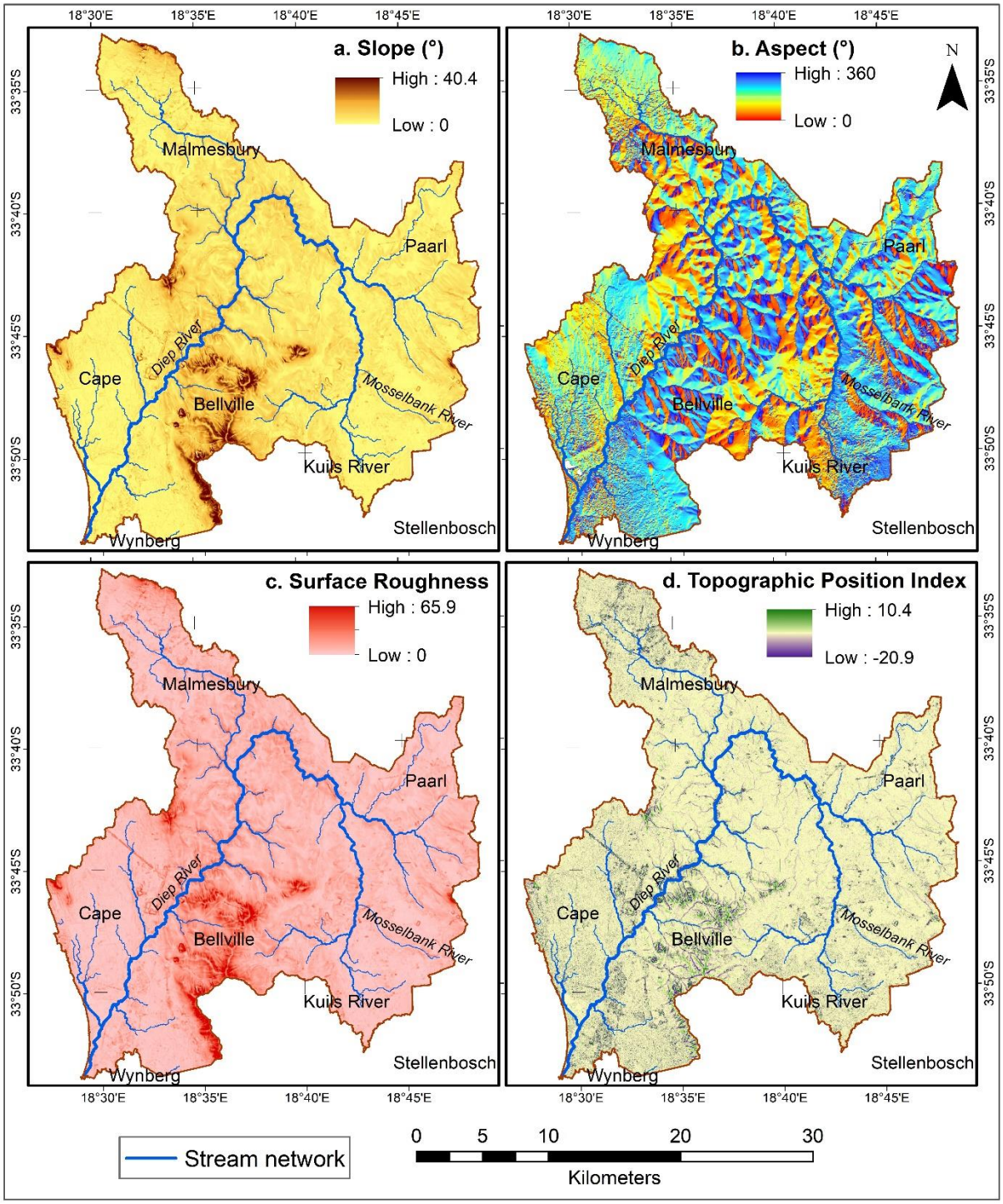


Figure 5.27: Maps of the terrain parameters, (a) slope (b) aspect (d) surface roughness, and (d) topographic position index

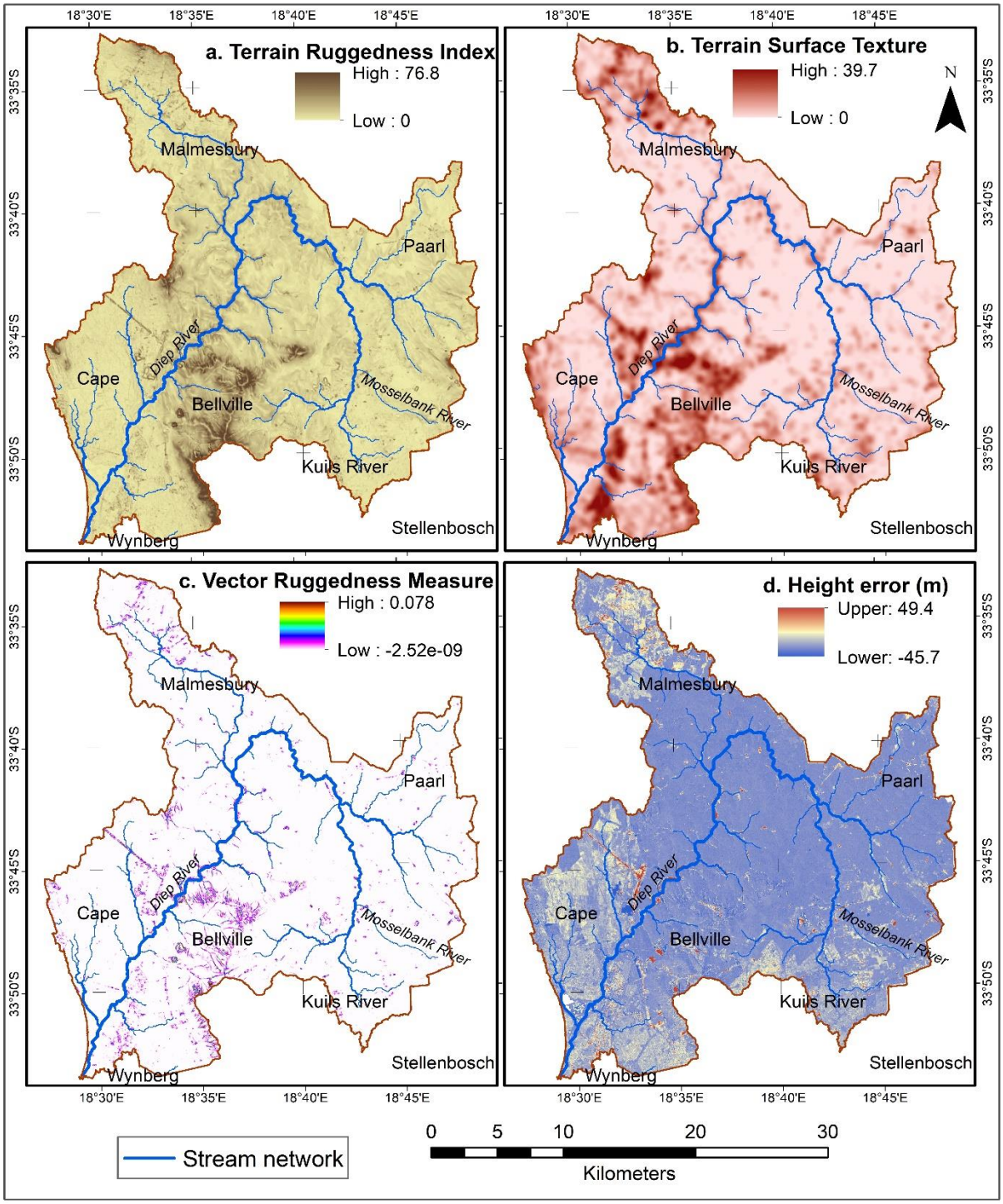


Figure 5.28: Maps of the terrain parameters – (a) terrain ruggedness index (b) terrain surface texture (c) vector ruggedness measure; and (d) height error map

Table 5.7: Descriptive statistics of the training/test data for evaluation of tree-based ensembles

Parameter	Min	25th percentile	50th percentile	75th percentile	Max	Mean	St Dev
Elevation (m)	16.52	49.68	62.37	76.35	154.31	64.63	21.66
Slope (°)	0.01	1.91	2.88	4.25	20.56	3.30	2.02
Aspect (°)	0.00	82.48	219.36	286.34	359.99	192.71	110.43
Surface Roughness	0.06	2.69	4.01	5.89	26.06	4.62	2.76
Topographic Position Index	-4.94	-0.16	0.04	0.21	8.26	0.00	0.61
Terrain Ruggedness Index	0.06	2.72	3.99	5.85	28.38	4.62	2.80
Terrain Surface Texture	0.00	0.67	2.07	5.45	24.94	3.92	4.62
Vector Ruggedness Measure	0.00	0.00	0.00	0.0002	0.0314	0.0003	0.0009

Note: 80% of the samples were used for training, while 20% were reserved for testing.

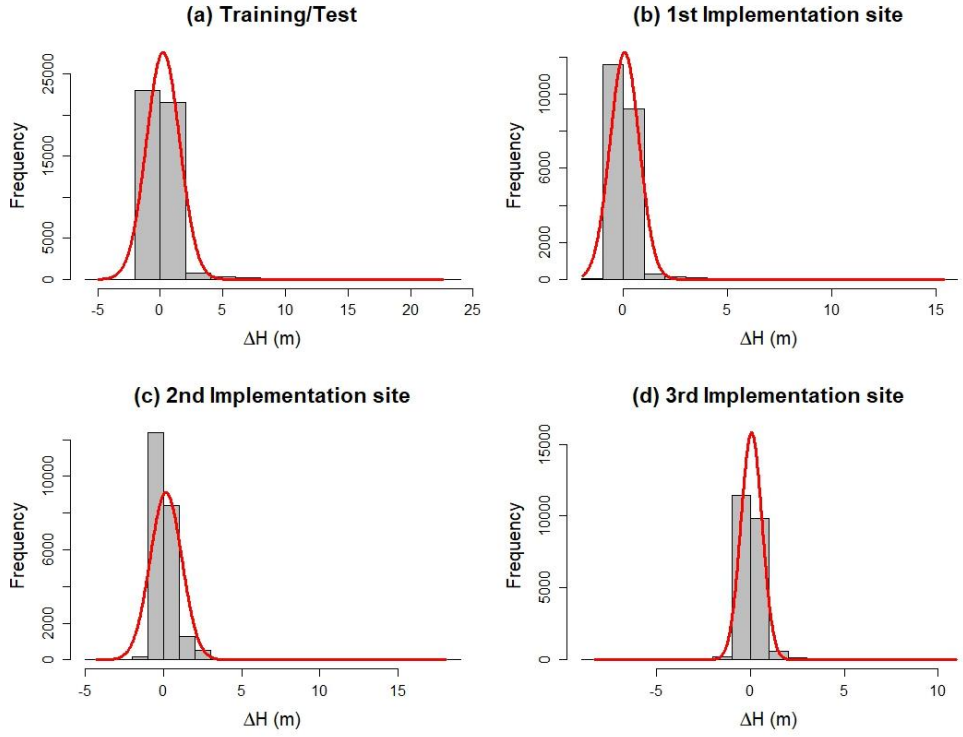


Figure 5.29: Histograms showing the elevation error (ΔH) distribution calculated from the full datasets, at the selected sites

5.3.2 Model performance indicators

Table 5.8 presents the test error, for a comparison of the performance of the models on the unseen test data. A combination of metrics is often required for assessing model performance. CatBoost emerged as the model with the lowest test MAE (0.310 m) and RMSE (0.704 m). The low test MdAE of several models, e.g., random forest (0.151 m), CatBoost (0.153 m), XGBoost (0.154 m) and LightGBM (0.155 m) suggests their robustness for modelling non-normally distributed data points.

Table 5.8: Test error based on the model default parameters. The lowest test errors are highlighted

Model	**MAE (m)**	**RMSE (m)**	**MdAE (m)**
Random Forest	0.312	0.731	**0.151**
Bagging Meta-estimator	0.327	0.757	0.160
AdaBoost	0.526	1.042	0.278
GBM	0.329	0.749	0.161
XGBoost	0.319	0.735	0.154
LightGBM	0.314	0.715	0.155
Histogram-based GBM	0.317	0.724	0.157
CatBoost	**0.310**	**0.704**	0.153
NGBoost	0.332	0.754	0.162
Regularised Greedy Forest	0.338	0.776	0.162

The model prediction errors are compared in Figure 5.30 and Figure 5.31 respectively.

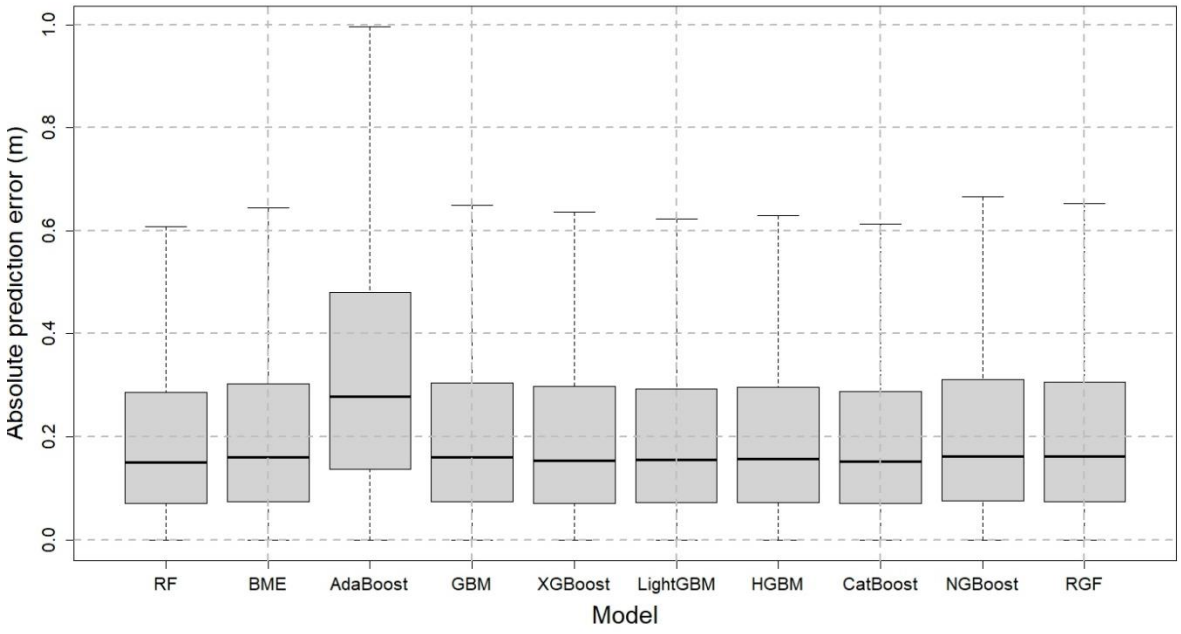


Figure 5.30: Comparison of model prediction error

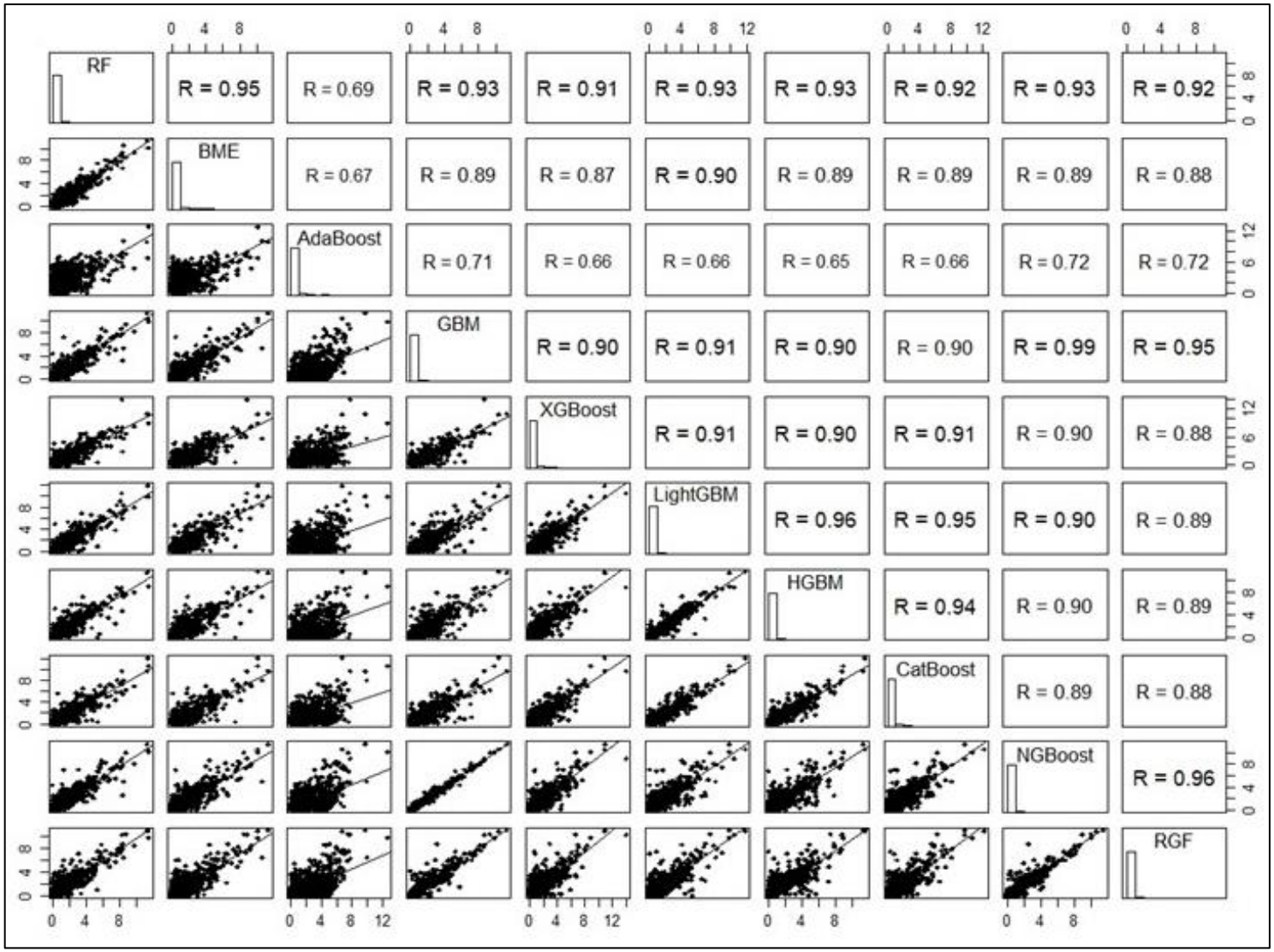


Figure 5.31: The absolute prediction error represented by histograms (diagonal panel), comparative scatter plots (lower panel) and correlations (upper panel)

Computational efficiency is a key factor used in evaluating model performance. Table 5.9 shows a comparison of the training and test execution times of the models. LightGBM had the fastest training time (0.318 s) than the other investigated methods. Several experiments on public datasets have also shown that LightGBM speeds up the training process of conventional GBDTs while achieving competitive accuracy (Ke et al., 2017). The histogram-based GBM recorded the second fastest training time. The histogram-based GBM was inspired by LightGBM and is reputed to be orders of magnitude faster than the GBM for large datasets. Moreover, it reduces the training time but without losing the accuracy level (Padhi et al., 2021). Despite their predictive accuracy performances, random forest and the RGF performed slower than most of the GBDTs. The slowest training time of 156.492 s was recorded by NGBoost. According to Duan et al. (2020), NGBoost requires far less expertise than other competing methods and it scales to large numbers of features with the same complexity as traditional boosters.

Despite giving competitive test accuracy, the GBM regressor could not match the computation time of the newer gradient-boosted trees (e.g., XGBoost, LightGBM and CatBoost). Random forest consumed a significantly higher training time of 49.7 s. This indicates that the newer GBDTs such as XGBoost, LightGBM and CatBoost could offer highly efficient computation times, especially for larger datasets.

Table 5.9: Comparison of execution time (in seconds) for the training and test sets. The fastest execution times are highlighted

Models	**Training time**	**Testing time**
Random Forest	49.697	0.504
Bagging Meta-estimator	4.850	0.049
AdaBoost	1.814	0.026
GBM	13.777	**0.021**
XGBoost	2.937	0.031
LightGBM	**0.318**	0.050
Histogram-based GBM	0.671	0.074
CatBoost	8.028	0.035
NGBoost	156.492	0.636
Regularised Greedy Forest	12.869	0.179

5.3.3 Accuracy assessment at implementation sites

Table 5.10 and Table 5.11 show the accuracy of the corrected DEMs at the implementation sites. At the 1st site, the corrections led to a 6 – 12% reduction in MAE and a 15 – 28% reduction in RMSE of the original DEM. Using the RMSE as a measure of vertical accuracy, Table 5.10 shows that the best correction was achieved by GBM (RMSE: 0.509 m) followed by the gradient boosting algorithms, XGBoost and NGBoost which both had RMSEs of 0.511 m. However, RGF outperformed some of the other gradient boosting algorithms such as LightGBM and CatBoost. The correction from AdaBoost was the least accurate (RMSE: 0.596 m), and there was a 28% increase in its MAE from 0.292 m to 0.372 m.

At the 2nd site, all the models achieved an improvement in accuracy shown by a 10 – 13% reduction in the MAEs and a 24 – 29% reduction in the RMSEs of the corrected DEMs. The exception is AdaBoost where the MAE increased by 6% from 0.429 m to 0.454 m. The best correction to the original DEM was attained by the two bagging models, random forest (RMSE: 0.735 m) and the Bagging Meta-estimator (RMSE: 0.744 m), followed by the gradient boosted algorithms. The least performing model was also AdaBoost (RMSE: 0.768 m). The accuracy achieved by XGBoost was closely matched by RGF.

Table 5.10: Accuracy measures of the corrected DEMs at the 1st and 2nd implementation sites. The best accuracies are highlighted

Tree-based Model	**1st implementation site**				**2nd implementation site**			
	Original Cop DEM		**Corrected Cop DEM**		**Original Cop DEM**		**Corrected Cop DEM**	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Random Forest	0.292	0.703	0.260	0.517	0.429	1.024	0.374	**0.735**
Bagging Meta-estimator			0.274	0.534			0.385	0.744
AdaBoost			0.372	0.596			0.454	0.768
GBM			0.261	**0.509**			0.376	0.757
XGBoost			0.261	0.511			0.377	0.758
LightGBM			0.258	0.519			**0.372**	0.746
Histogram-based GBM			**0.258**	0.519			0.373	0.752
CatBoost			0.258	0.522			0.374	0.756
NGBoost			0.262	0.511			0.377	0.757
Regularised Greedy Forest			0.263	0.517			0.377	0.758

Table 5.11: Accuracy measures of the corrected DEMs at 3rd implementation site, and the whole area. The best accuracies are highlighted

Tree-based Model	**3rd implementation site**				**Whole area (1st site + 2nd site + 3rd site)**			
	Original Cop DEM		**Corrected Cop DEM**		**Original Cop DEM**		**Corrected Cop DEM**	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Random Forest	0.294	0.567	0.296	0.636	0.340	0.792	0.311	0.638
Bagging Meta-estimator			0.305	0.657			0.322	0.653
AdaBoost			0.388	0.680			0.405	0.687
GBM			0.290	0.617			0.311	0.639
XGBoost			**0.289**	0.614			0.311	0.638
LightGBM			0.290	0.613			**0.308**	**0.635**
HGBM			0.291	0.618			0.309	0.640
CatBoost			0.294	0.627			0.310	0.645
NGBoost			**0.289**	0.612			0.311	0.637
Regularised Greedy Forest			0.290	**0.609**			0.312	0.639

Unexpectedly, there were little or no accuracy improvements at the 3rd site. Although there was a reduction in the MAE in most cases, the RMSE of the original DEM escalated. The reason for the decline in accuracy is unclear, especially because all selected sites have similar terrain characteristics. Despite the outcome, RGF achieved the lowest RMSE of 0.609 m, while XGBoost and NGBoost achieved the lowest MAE of 0.289 m. Several innovations were introduced in RGF (e.g., tree-structured regularisation and a fully-corrective regularised greedy algorithm) to improve its robustness for learning nonlinear decision forests (Johnson & Zhang, 2014). The accuracy achieved in the whole area indicates that the best model performance is achieved when the trained models are deployed at targeted sites with similar terrain characteristics with the training data, and that generalisation could degrade the performance.

Figure 5.32 shows a visual comparison of the height error maps, calculated from the original and corrected DEMs while Figure 5.33 shows a visual comparison of the original and corrected DEMs at the three implementation sites. Generally, the corrected DEMs manifest several topographic improvements, refinement of the land surface and diminution of elevation bias. This proves the capability of tree-based ensembles for the enhancement of global DEMs.

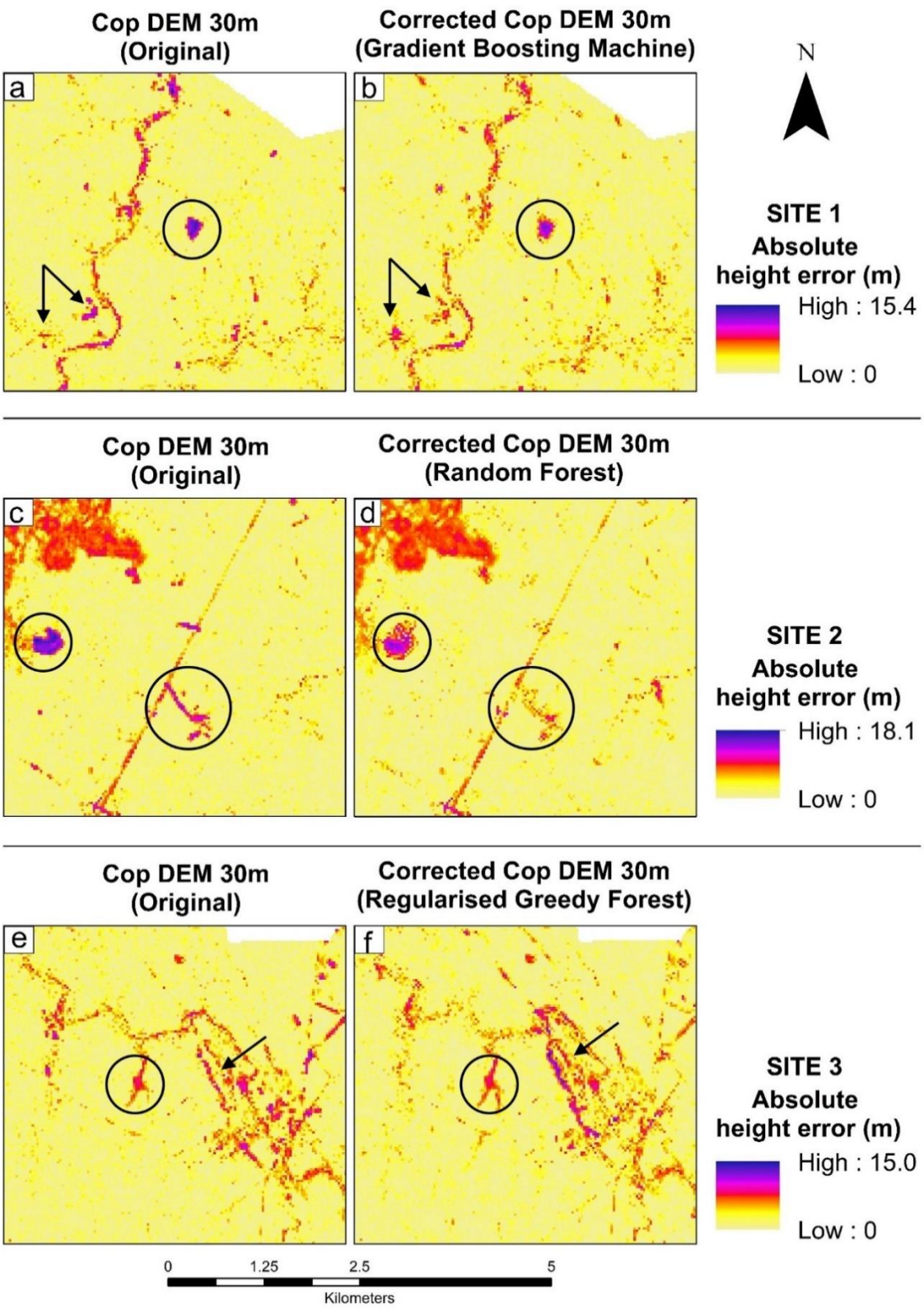


Figure 5.32: Visual comparison of the height error maps, calculated from the original and corrected DEMs

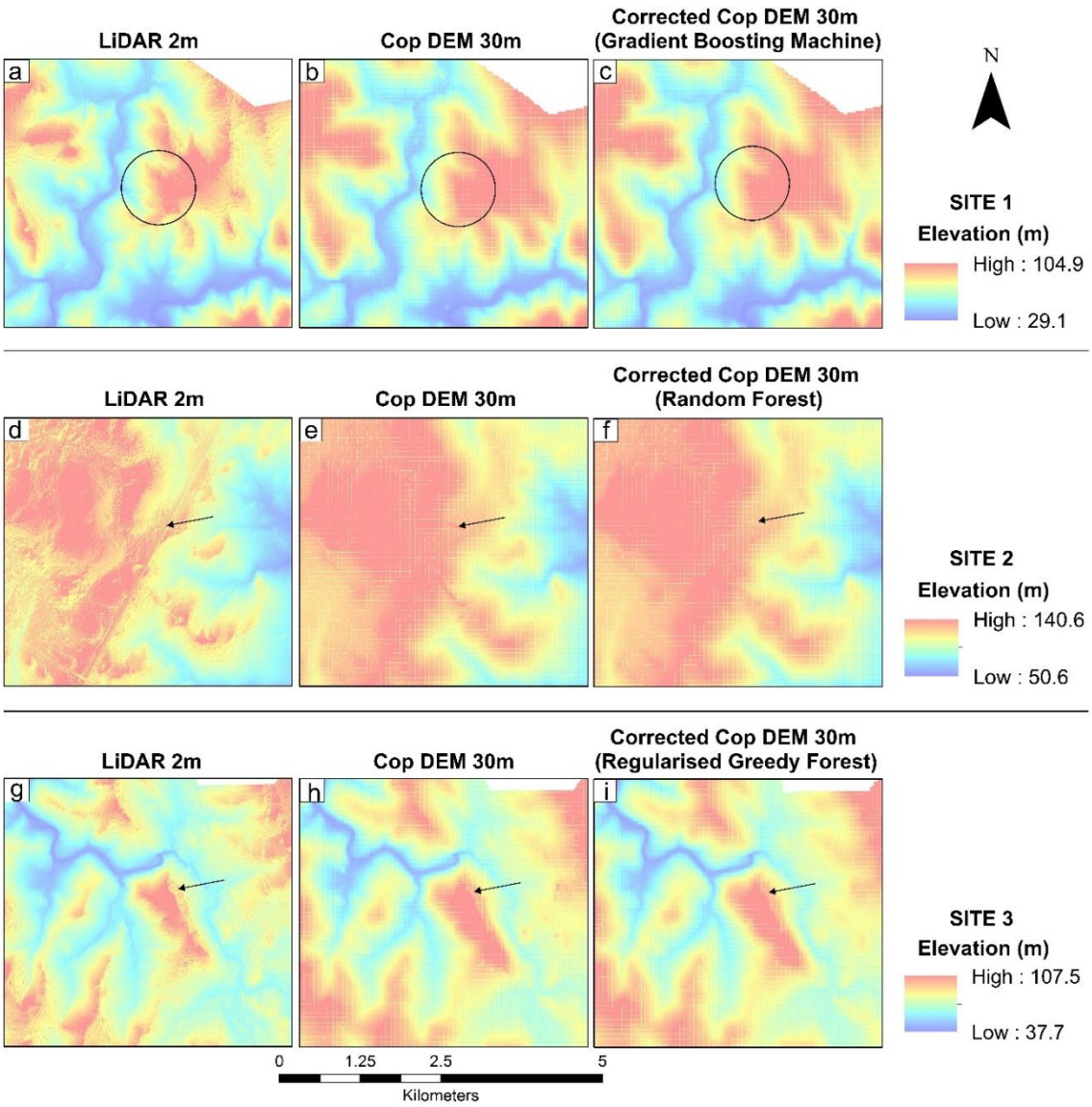


Figure 5.33: Visual comparison of the original and corrected DEMs at the three implementation sites

5.3.4 Explainability of tree-based ensembles

The general notion is that machine learning algorithms are a 'black box', in other words, their internal working mechanisms and how they generate the predictions are most often not comprehensible. This is not absolutely true, especially for the tree-based models. In this section, feature importance plots and partial dependence plots (PDPs) are used to explain how the models arrive at certain results. This in turn, makes it easier to understand why the models generate certain results and the deriving factors. Figure 5.34 shows the feature importance plots for some selected models to provide insight on the relevance of the features in the construction of the tree-based models. The first-order terrain derivatives, slope and aspect, demonstrate moderate or minimal influence in the prediction of the elevation error. The TPI, TST and VRM emerged as the most influential predictors. This sensitivity of the VRM is possibly due to its

ability to incorporate 3D dispersion of vectors and its ability to differentiate different types of terrain. Similarly, with the TPI, different landform types can be differentiated. RGF was unable to emphasize the importance of most features, as observed in the very low importance allocated to elevation, slope, aspect, surface roughness and TRI. The influence of most features is better emphasized by CatBoost and LightGBM. The TPI, TST and VRM are generally the most influential terrain parameters for DEM error prediction by random forest, XGBoost CatBoost and RGF. The GBM exploits to a greater degree the influences of TPI, VRM and TST but other feature such as slope, surface roughness and TRI are less influential. Figure 5.35 shows the feature importance of the NGBoost-derived location and scale parameters, based on the assumption that the feature values are normally distributed. For normally distributed data, the location and scale parameters correspond to the mean and standard deviation, respectively (Stanford ML Group, 2023).

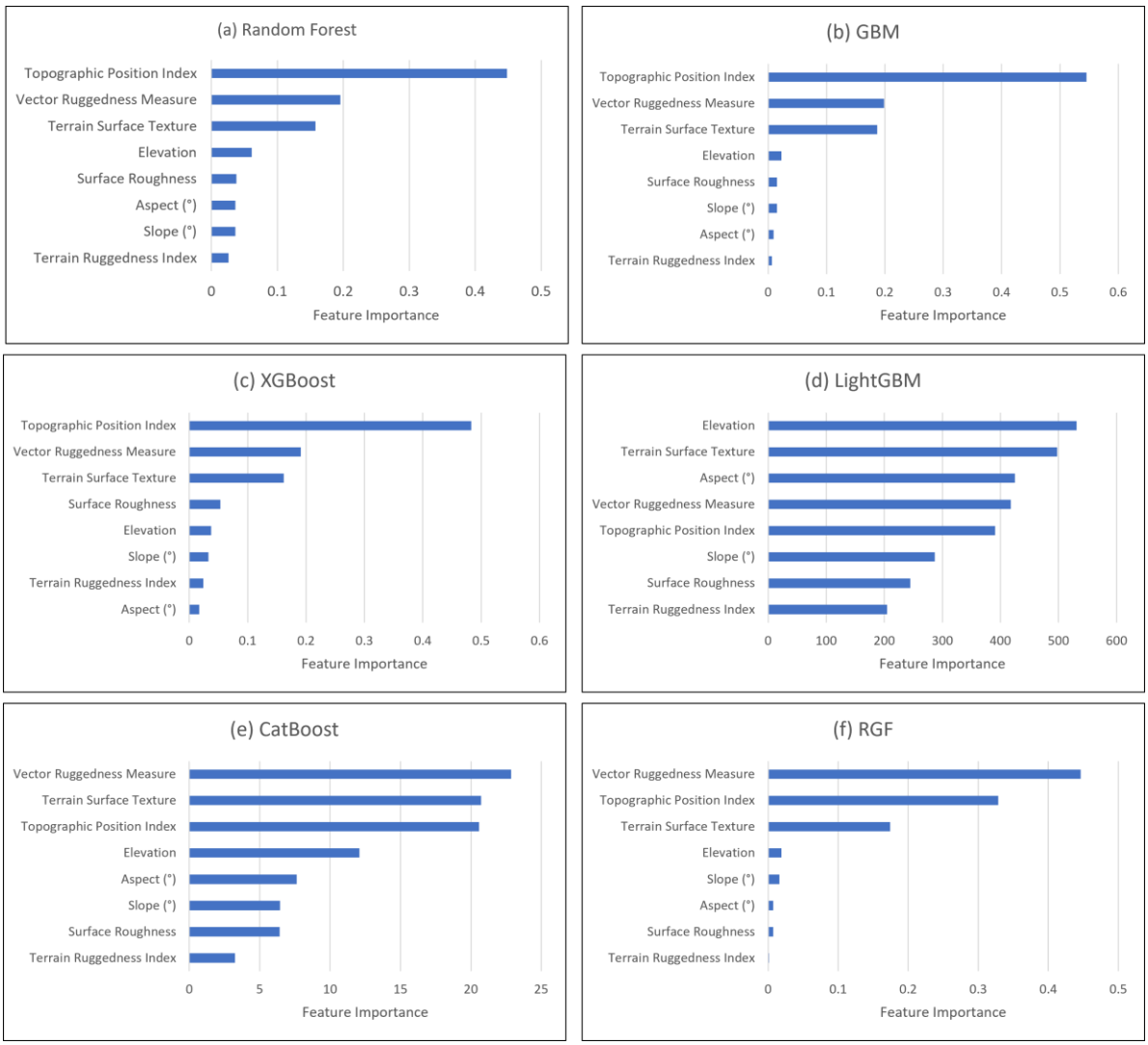


Figure 5.34: Feature importance plots shown for selected models – (a) Random forest (b) GBM (c) XGBoost (d) LightGBM (e) CatBoost (f) RGF

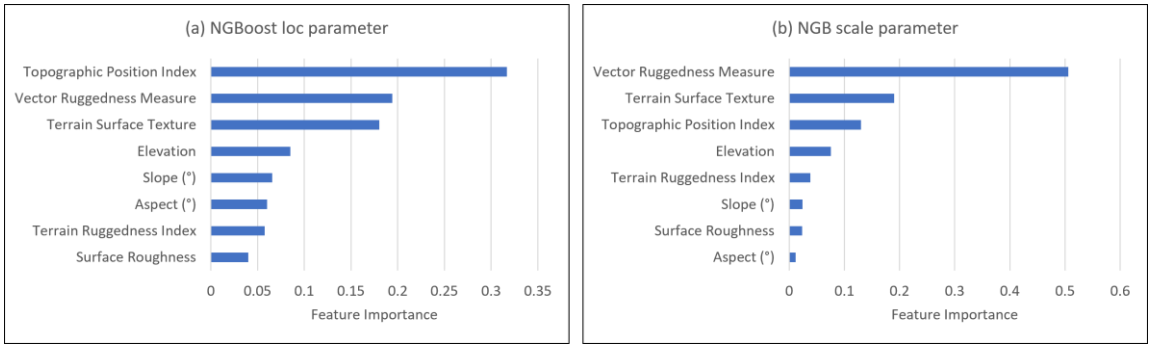


Figure 5.35: NGBoost feature importance plots for distribution parameters – (a) location parameter (b) scale parameter

PDPs depict the functional relationship between a small number of input variables and predictions. It can show how the predictions partially depend on values of the input variables of interest. Figure 5.36 - Figure 5.38 show the PDP plots for GBM, random forest and RGF respectively. The plots depict the functional relationship and interaction between the input variables (e.g., elevation, slope, aspect) and predicted elevation errors. The trend line captures how the predicted errors change as the feature values change. The PDPs corroborate the observation from the feature importance plots which showed that TPI, TST and VRM are influential predictors. The most significant increases in the errors are attributable to the higher values of TPI, VRM and TST. In random forest and GBM, the predicted errors are almost indifferent to variations in slope, surface roughness and TRI.

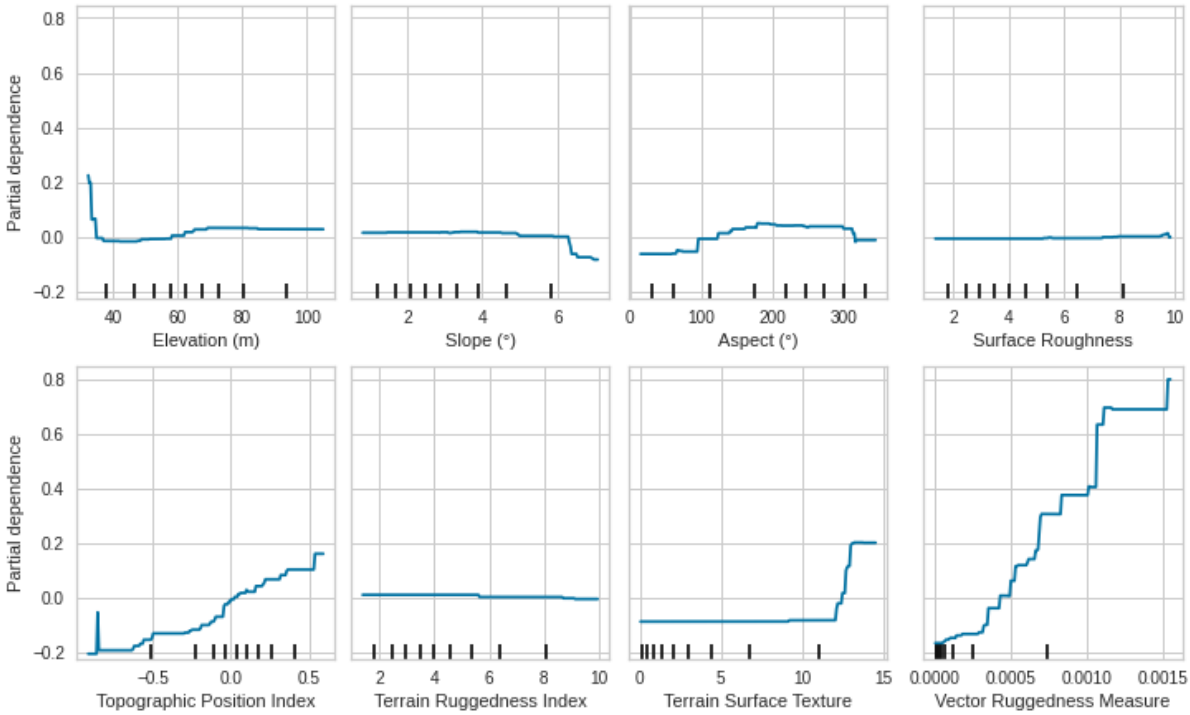


Figure 5.36: Partial dependence plot of the best performing model (lowest RMSE) at the 1st implementation site– GBM. The y-axis shows the partial dependence while the x-axis shows the feature class values

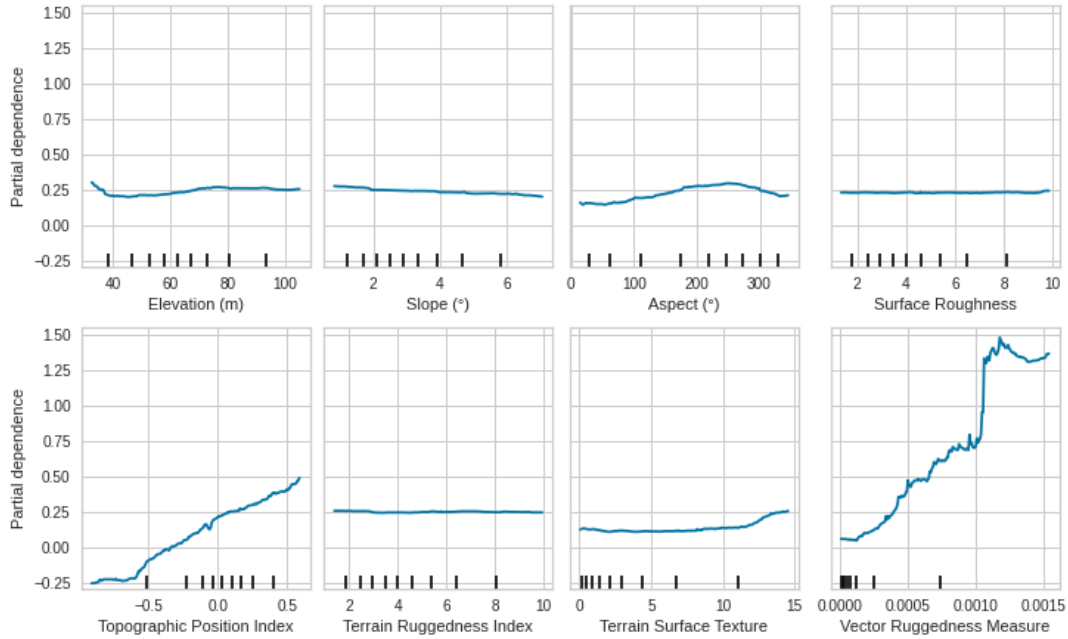


Figure 5.37: Partial dependence plot of the best performing model (lowest RMSE) at the 2nd implementation site – random forest. The y-axis shows the partial dependence while the x-axis shows the feature class values.

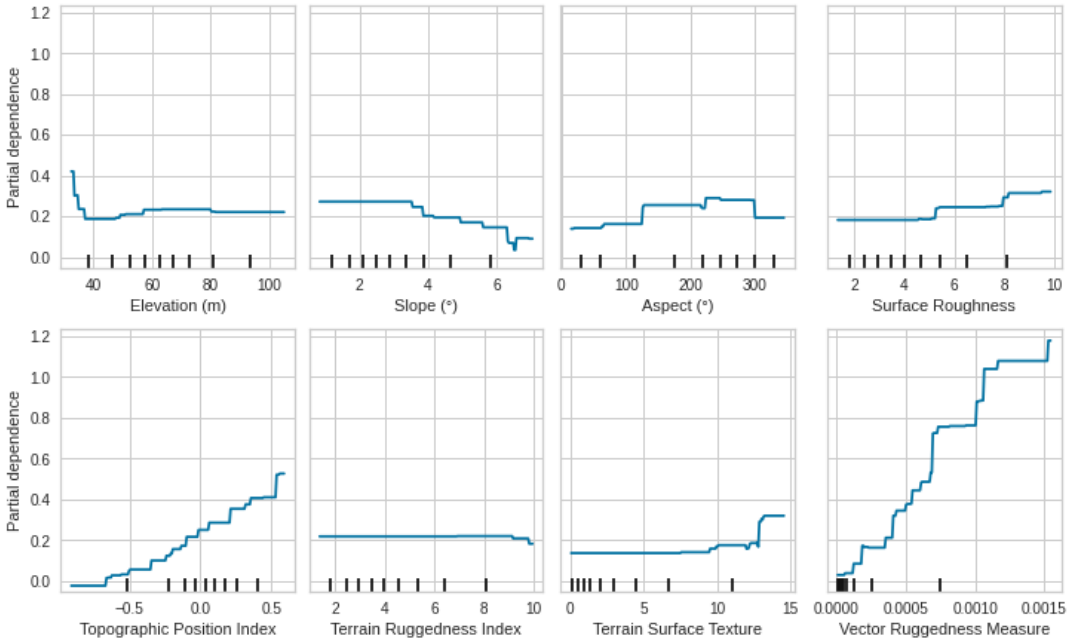


Figure 5.38: Partial dependence plot of the best performing model (lowest RMSE) at the 3rd implementation site – regularised greedy forest. The y-axis shows the partial dependence while the x-axis shows the feature class values.

5.3.5 Discussion of tree-based ensemble DEM enhancement results

The results demonstrate the capability of tree-based ensembles for enhancing the quality of digital elevation models. The consensus from the evaluated models indicates that TPI, TST and VRM are the most influential terrain parameters for predicting the elevation error in agricultural landscapes. Slope, aspect, surface roughness and TRI are also important predictors

but had the least influence in the study area. The NGBoost algorithm, an innovative probabilistic regression algorithm which utilises natural gradients (instead of regular gradients) has been tested for returning point estimates of elevation error. Thus, it has been proven that natural gradients can deliver satisfactory results when applied to remote sensing tasks. The choice of algorithm could depend on the desired user requirements and available computing resources. For example, while NGBoost might require far less expertise, its learning task parameters have not been fully developed to the extent of the more robust GBDTs such as XGBoost. The comparison of bagging and boosting ensembles shows that both techniques are recommendable with respect to reducing the bias/variance in the prediction errors. This comparative analysis serves as an invaluable source of knowledge on the performances of tree-based ensembles for handling remote sensing tasks. It is possible to improve the predictive power and/or accuracy of the algorithms, e.g., through stacking whereby the algorithms are integrated via a meta-regressor, or through optimisation. However, since their predictions are already complementary and competitive, the optimisation option is favoured. Several factors favour the selection of XGBoost, LightGBM and CatBoost for the next stage which involves the development of a comprehensive DEM enhancement framework:

1. All three of the models (XGBoost, LightGBM and CatBoost) are very recent implementations of gradient boosting, an extremely powerful boosting algorithm which to the author's knowledge, is yet to be tested for DEM enhancement.
2. All three models provide state-of-the-art results and are competitive with leading machine learning and deep learning algorithms.
3. Unlike some other three-based models, XGBoost, LightGBM and CatBoost have extensive, comprehensive and excellent documentations of their training parameters, learning task parameters and hyperparameters. These documentations are very essential for the process of ablation, troubleshooting and the multistep process of debugging.
4. The accuracies achieved by XGBoost, LightGBM and CatBoost were very competitive in the comparative evaluation. For example, LightGBM and CatBoost rendered the highest accuracies at the 1st implementation site.
5. Finally, according to the No-Free-Lunch (NFL) theorem, it can never be concluded that some learning algorithms are more sensible than others, and even good learning algorithms have limitations (Goldblum et al., 2023; Sterkenburg & Grünwald, 2021). In the original interpretation by (Wolpert, 1995, 2002), "all learning algorithms are equivalent". There is no universal algorithm that can solve all types of problems.

Having established the competitiveness of XGBoost, LightGBM and CatBoost, attention is given to the optimisation of their performances.

5.4 GBDT-based DEM Enhancement Framework

Following the previous assessment, three gradient boosted trees, XGBoost, LightGBM and CatBoost were selected for full implementation and optimisation. Their excellent and comparable performance with the other tree-based models coupled with their very recent release justified their selection. Moreover, there is extensive documentation on the hyperparameters of these three models.

5.4.1 Terrain characterisation

Figure 5.39 - Figure 5.49 show maps of the elevation distribution and terrain parameters derived from LiDAR, Copernicus and AW3D DEMs, including the maps of land cover (e.g., building footprints) and terrain parameters. Cape Town is topographically and geomorphologically diverse. In Figure 5.39, the imposing mountain ranges along the extensive Cape Peninsula are visible. On the northern end of the peninsula is the imposing flat-topped Table Mountain and neighbouring peaks, while the southern end of the peninsula juts into the Atlantic Ocean. Other parts of Cape Town lie on the flats further inland. The south-west and northeast-facing steep slopes of the Cape Peninsula and Table Mountain (Figure 5.41) dictate the characteristics of river drainage systems in that vicinity. The terrain classification done with the Topographic Position Index (TPI) shown in Figure 5.43 can distinguish topographic features in Cape Town (e.g., the peak of Table Mountain, flat plains in the urban/industrial areas and agricultural lands, upper and lower slopes etc.). The TRI (Figure 5.44) quantifies the elevation difference between adjacent elevation pixels and this provides a detailed depiction of the jaggedness or flatness of the terrain. Tree cover (Figure 5.47) is considerably higher on the Table mountain, and in other areas such as Somerset West, Strand and the Lourensford dam. The tree cover is sparse in the urban and industrial districts in the flatlands. The loss of vegetation is a natural consequence of urbanisation and this explains the high percentage of bare ground in the flat lands. Continuous urban development is associated with construction activities and paving of road surfaces which could smoothen out the surface ruggedness and texture in urban areas. Such urban areas are generally at the lower end of the surface roughness, TRI and TST scale. The high terrain ruggedness and roughness observed up north in the Koeberg Nature Reserve is probably due to the intersperseness and unevenness of the grasses, shrubs and bare ground.

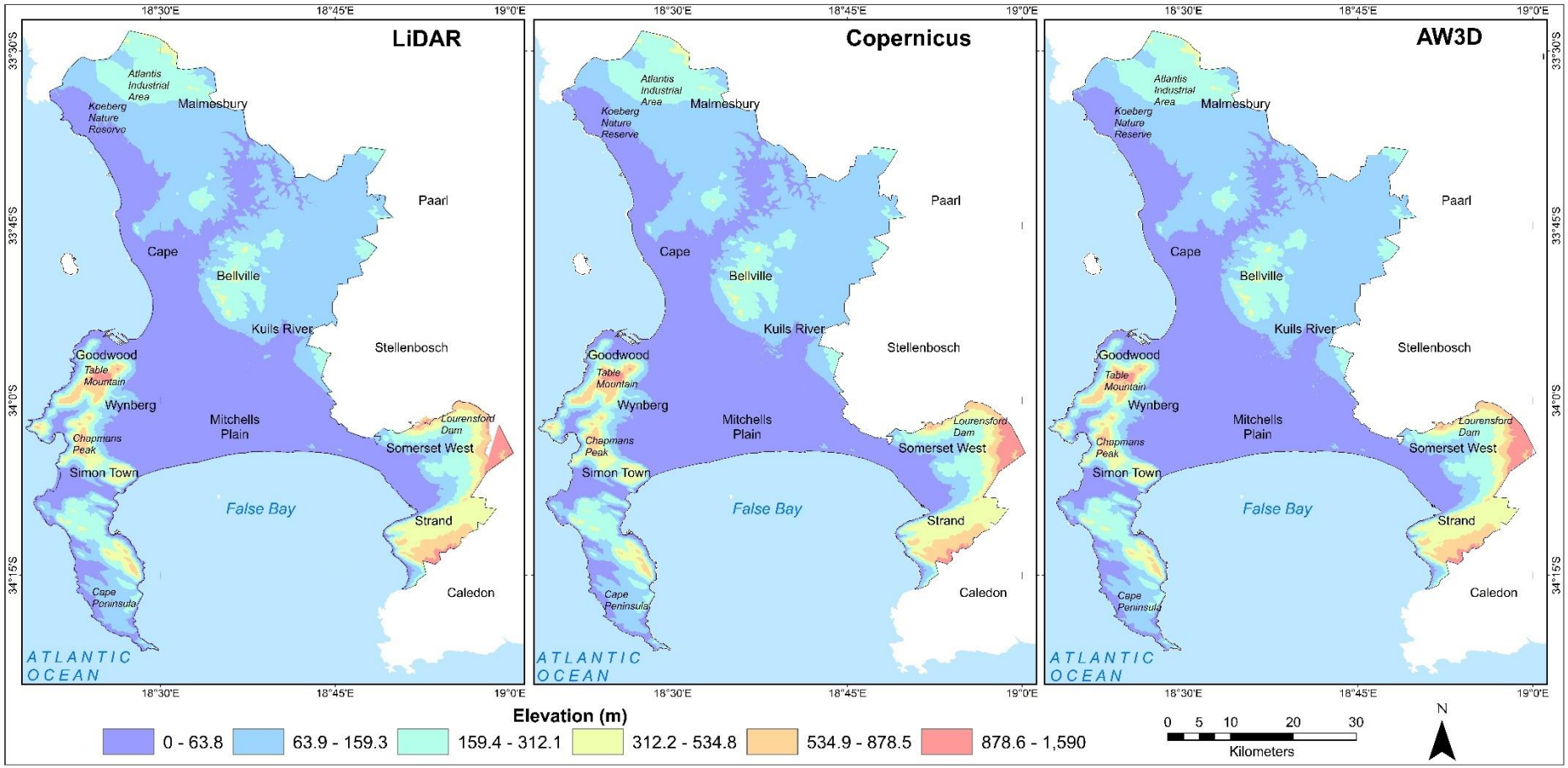


Figure 5.39: Visualisation of Cape Town topography with elevation data extracted from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

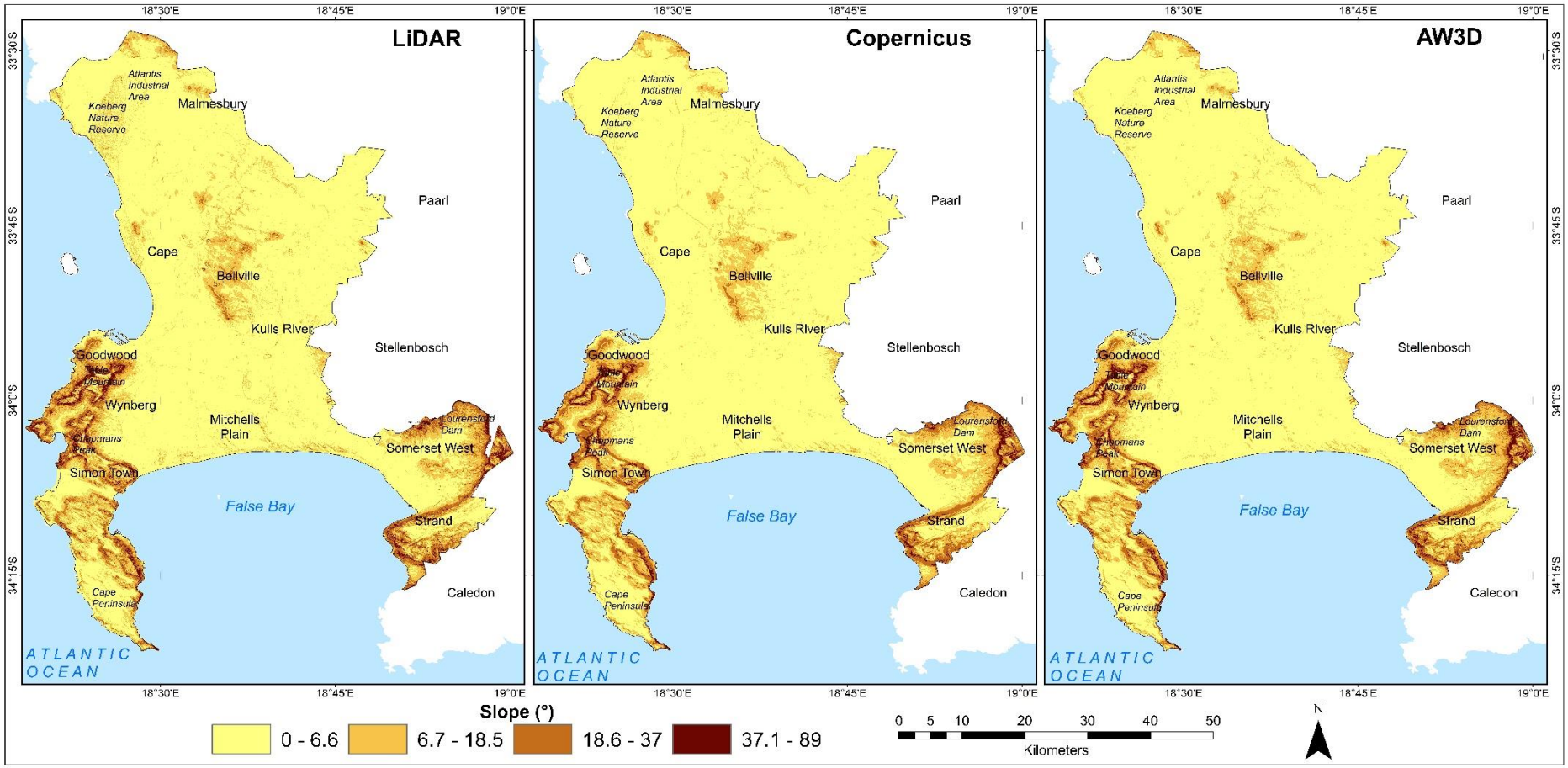


Figure 5.40: Visualisation of the rise and fall of the terrain in Cape Town using slope information extracted from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

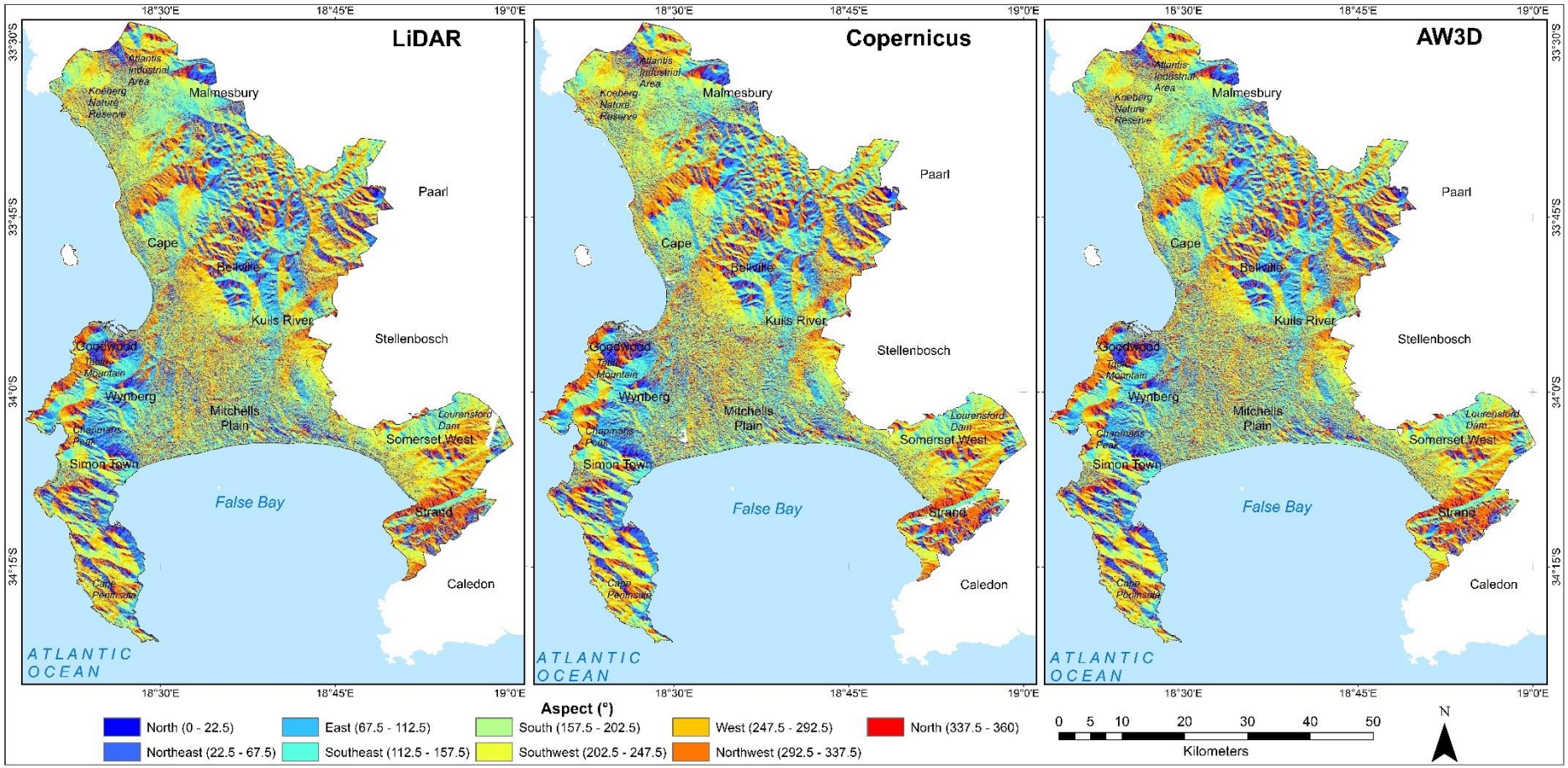


Figure 5.41: Colorised representation of aspect (slope direction) in Cape Town derived from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

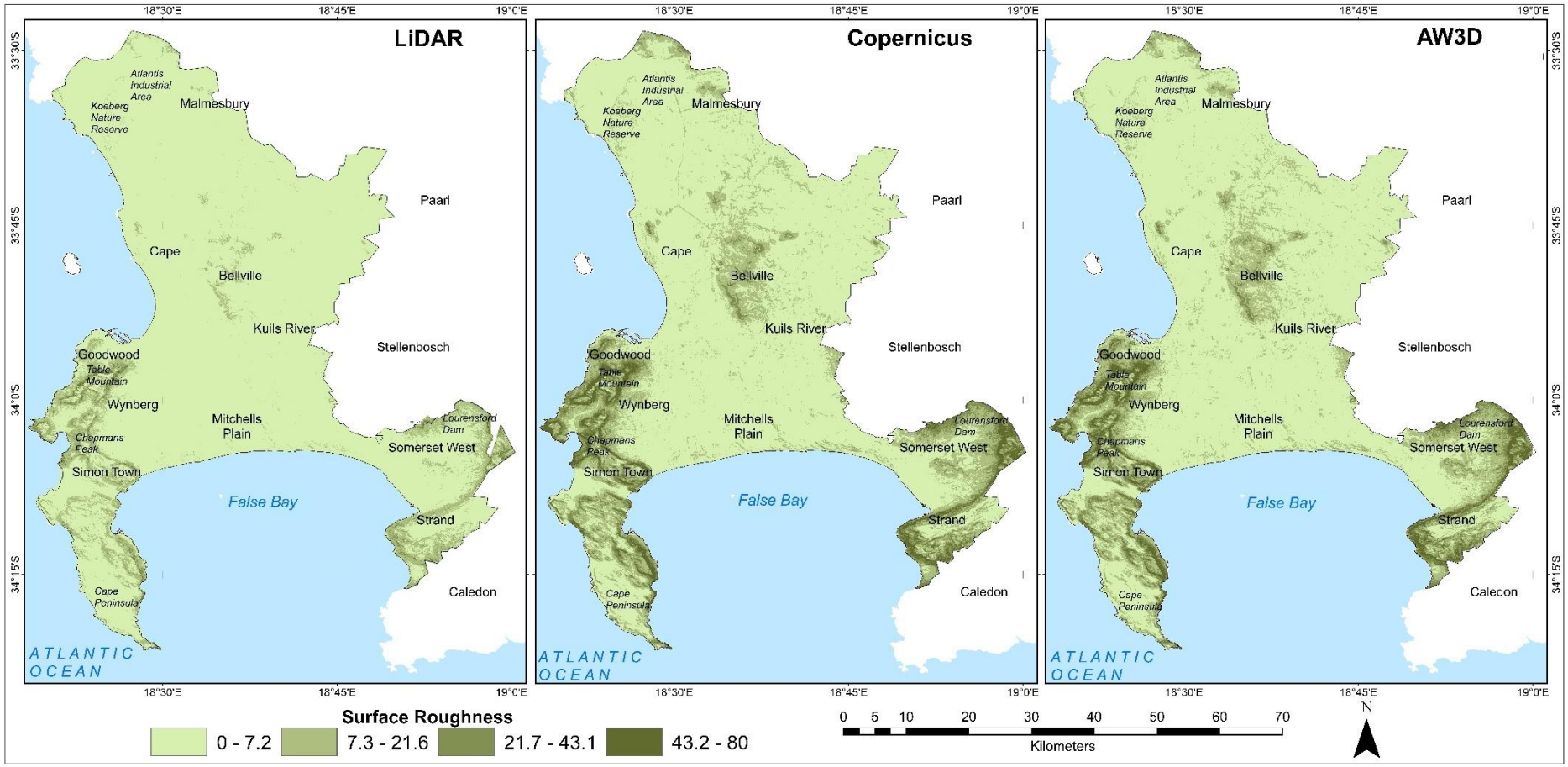


Figure 5.42: Depiction of the unevenness, variability and irregularity in the topography of Cape Town using surface roughness data derived from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

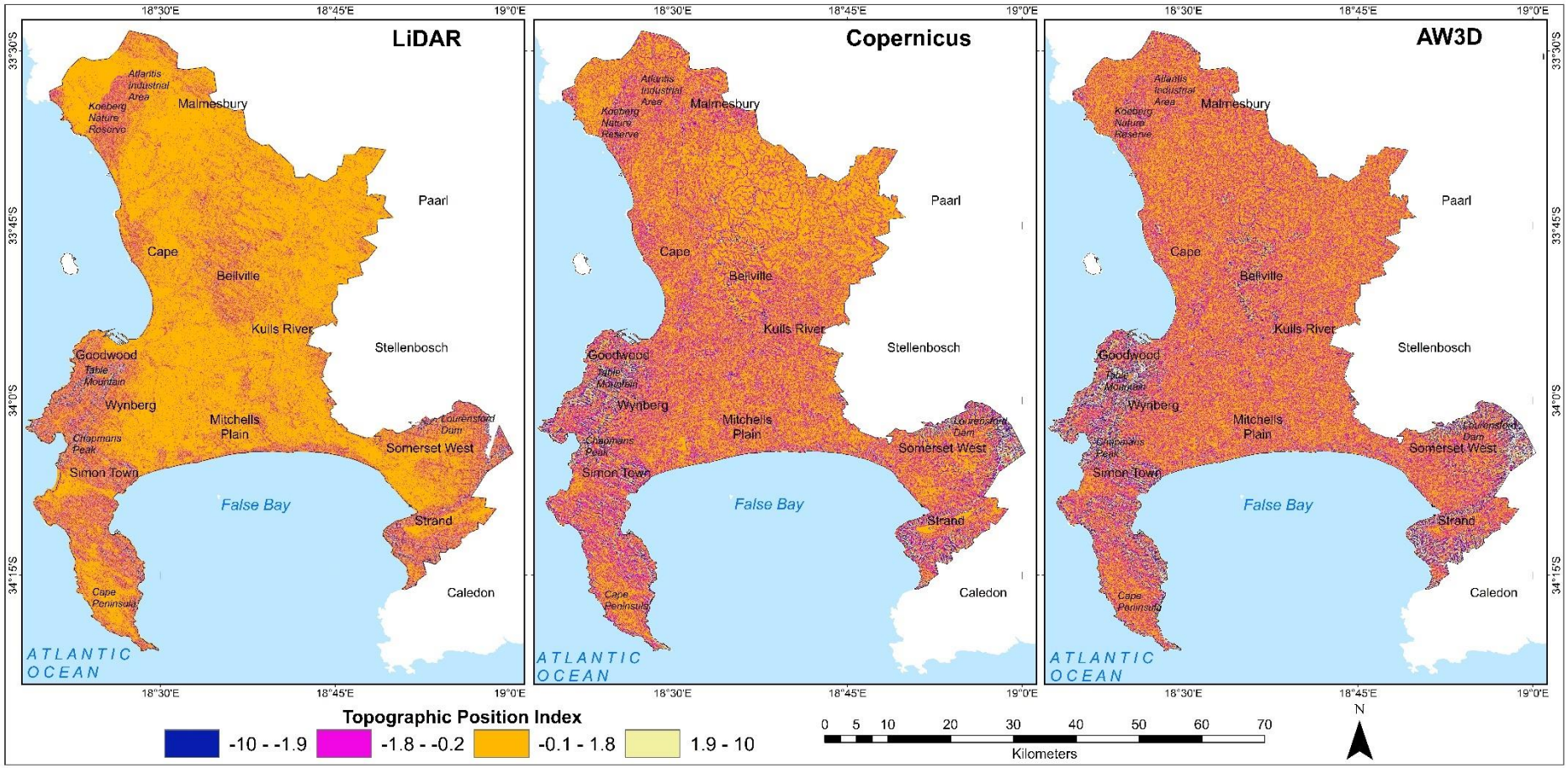


Figure 5.43: The TPI information shown is derived from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

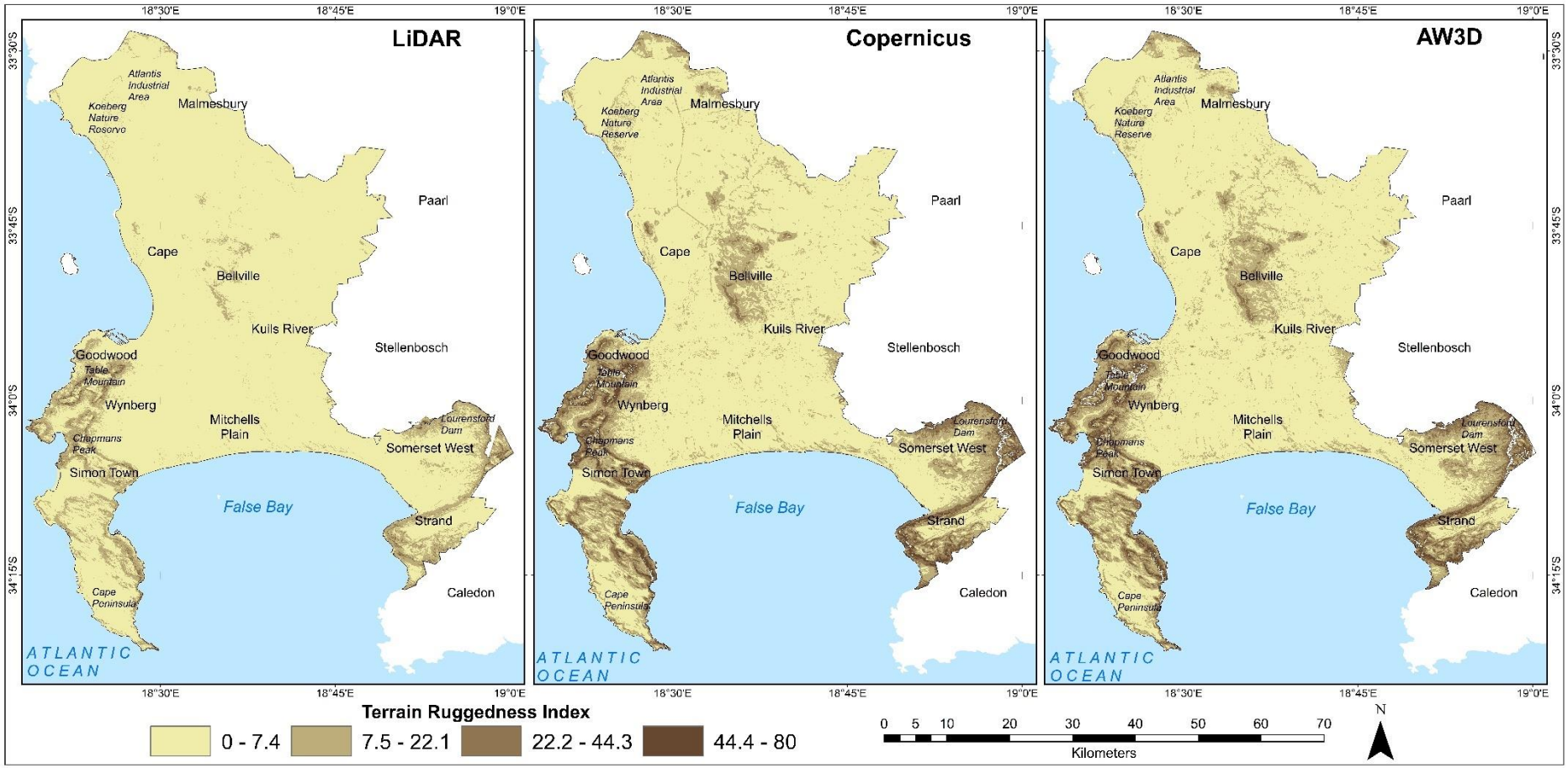


Figure 5.44: Visualisation of the topographic heterogeneity of Cape Town with the terrain ruggedness index (TRI) derived from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

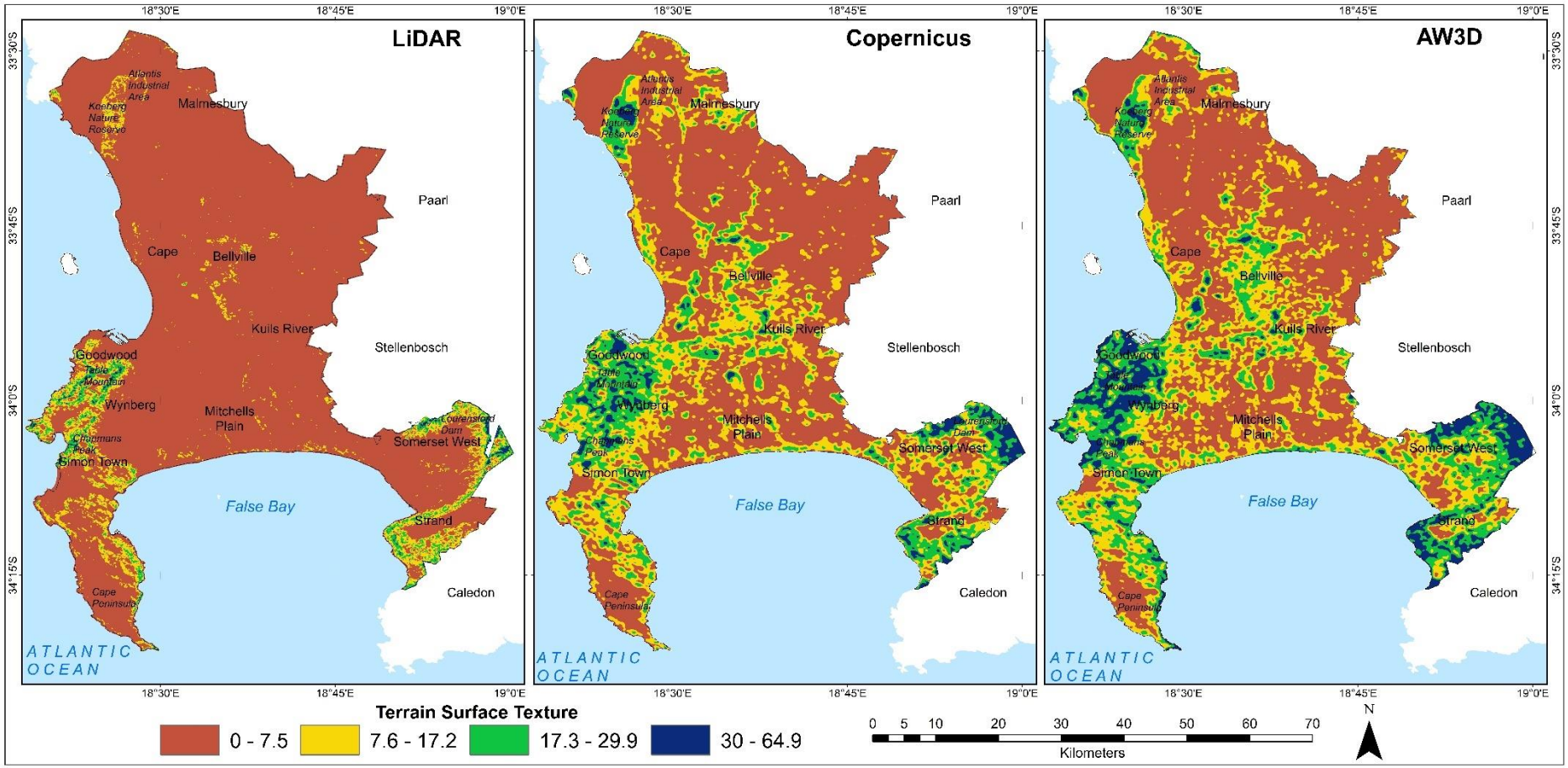


Figure 5.45: Colourised representation of the terrain surface texture of Cape Town, derived from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

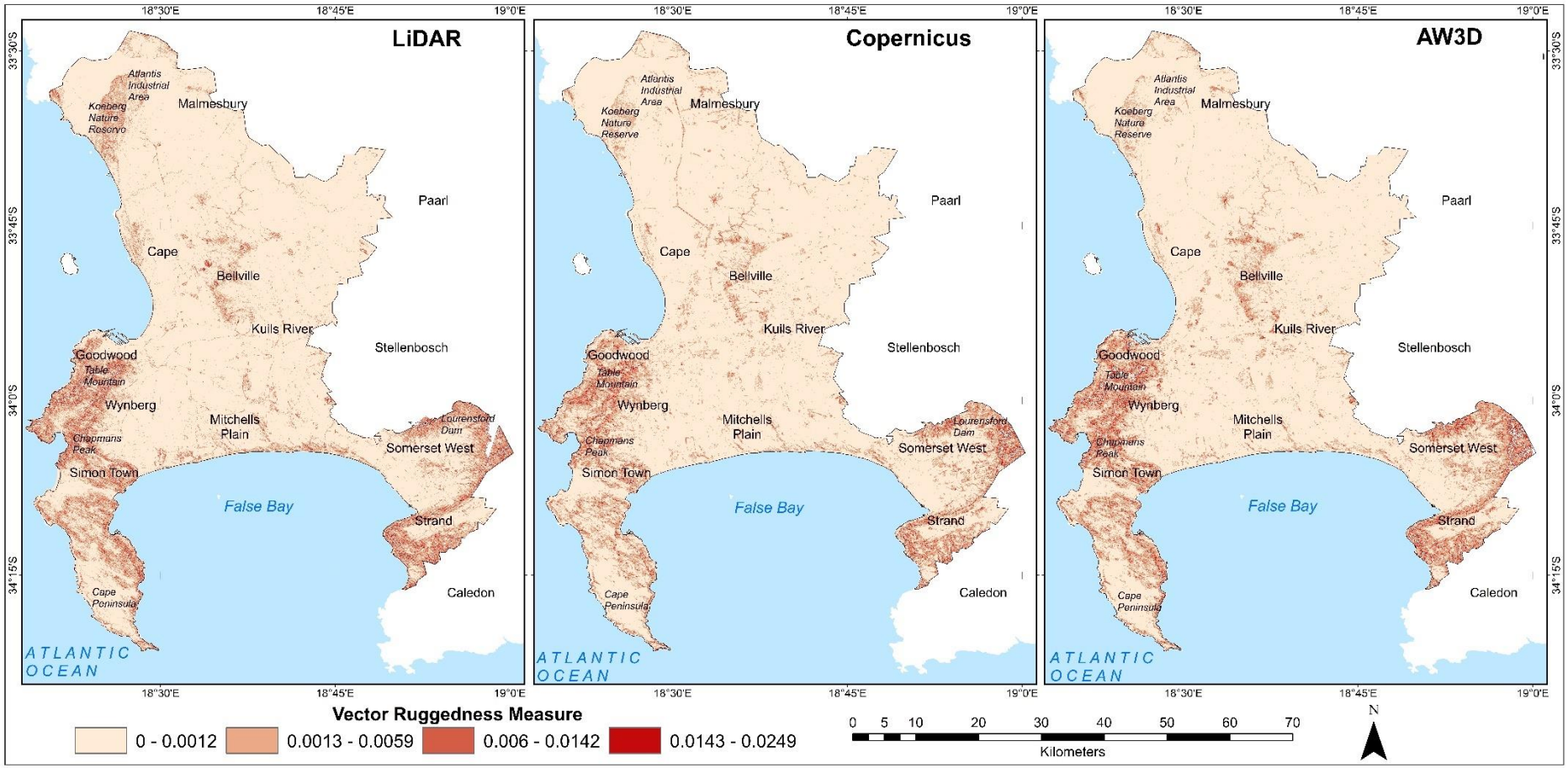


Figure 5.46: A different perspective of the ruggedness of Cape Town topography, using the vector ruggedness measure (VRM) derived from LiDAR (left), Copernicus DEM (middle) and AW3D DEM (right)

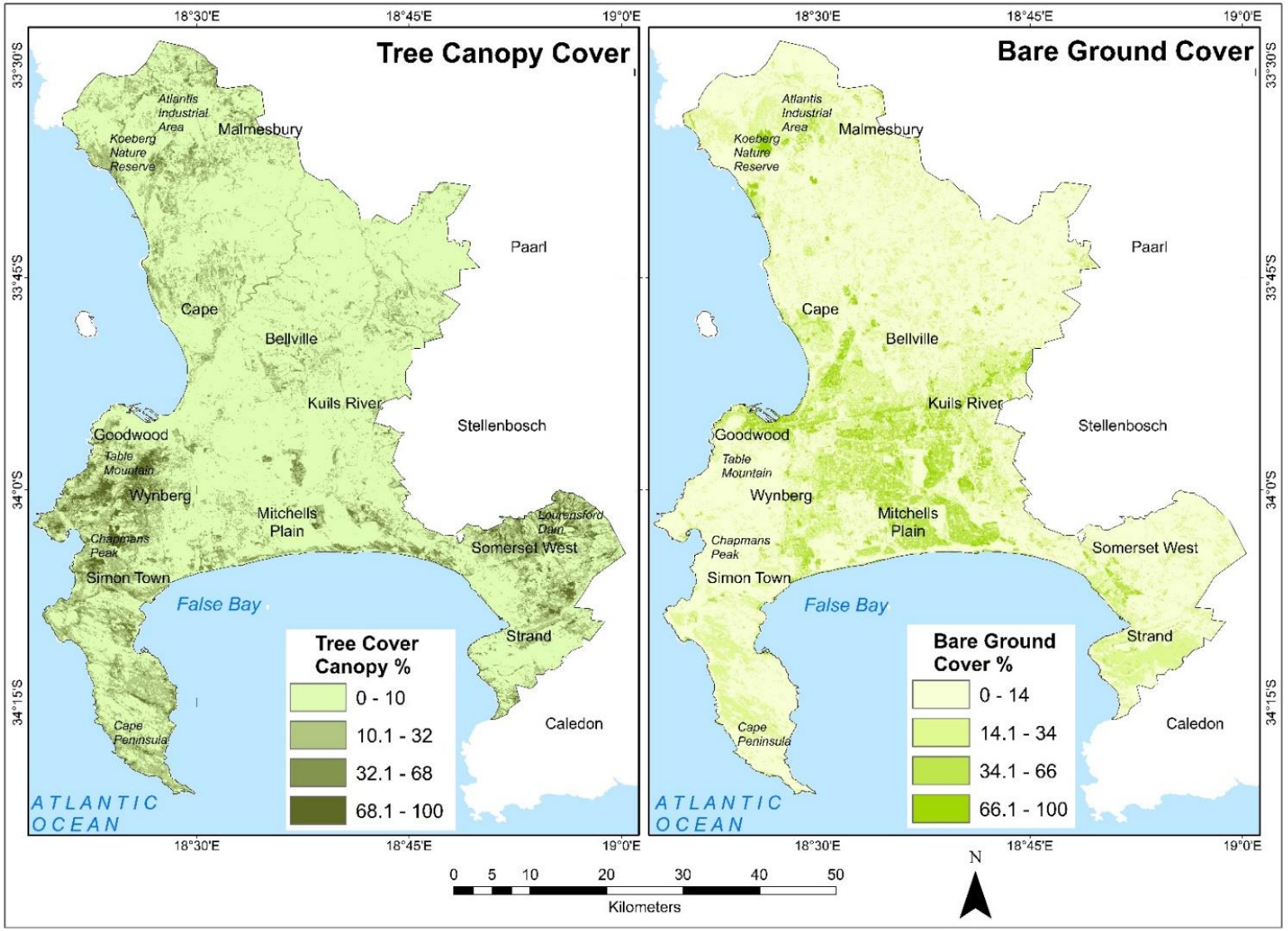


Figure 5.47: Representation of the percentage tree cover (left) and percentage bare ground cover (right) in Cape Town, derived from the Global Land Analysis and Discovery (GLAD) database

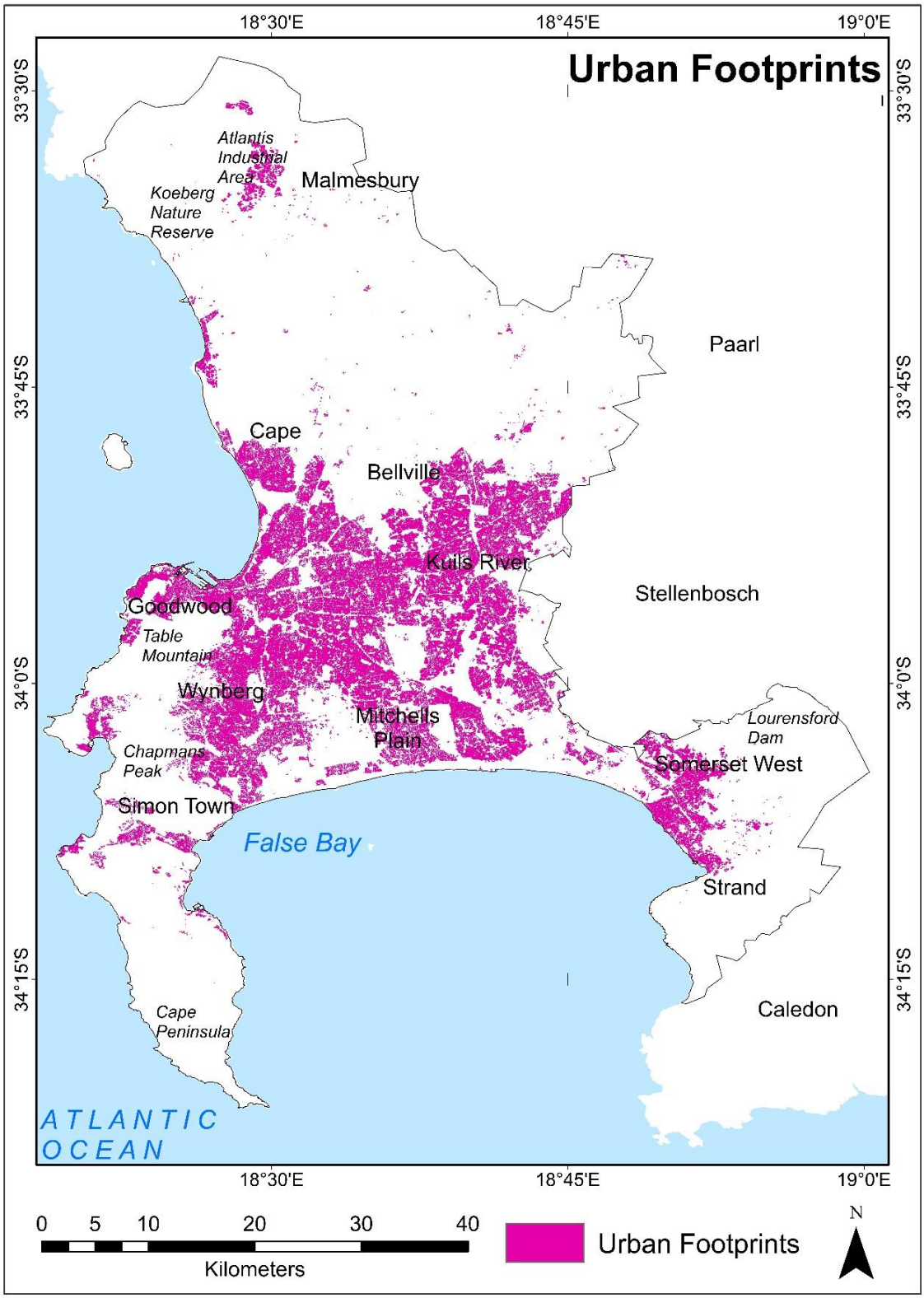


Figure 5.48: Spatial distribution of urban footprints in Cape Town derived from the DLR Global urban footprints dataset

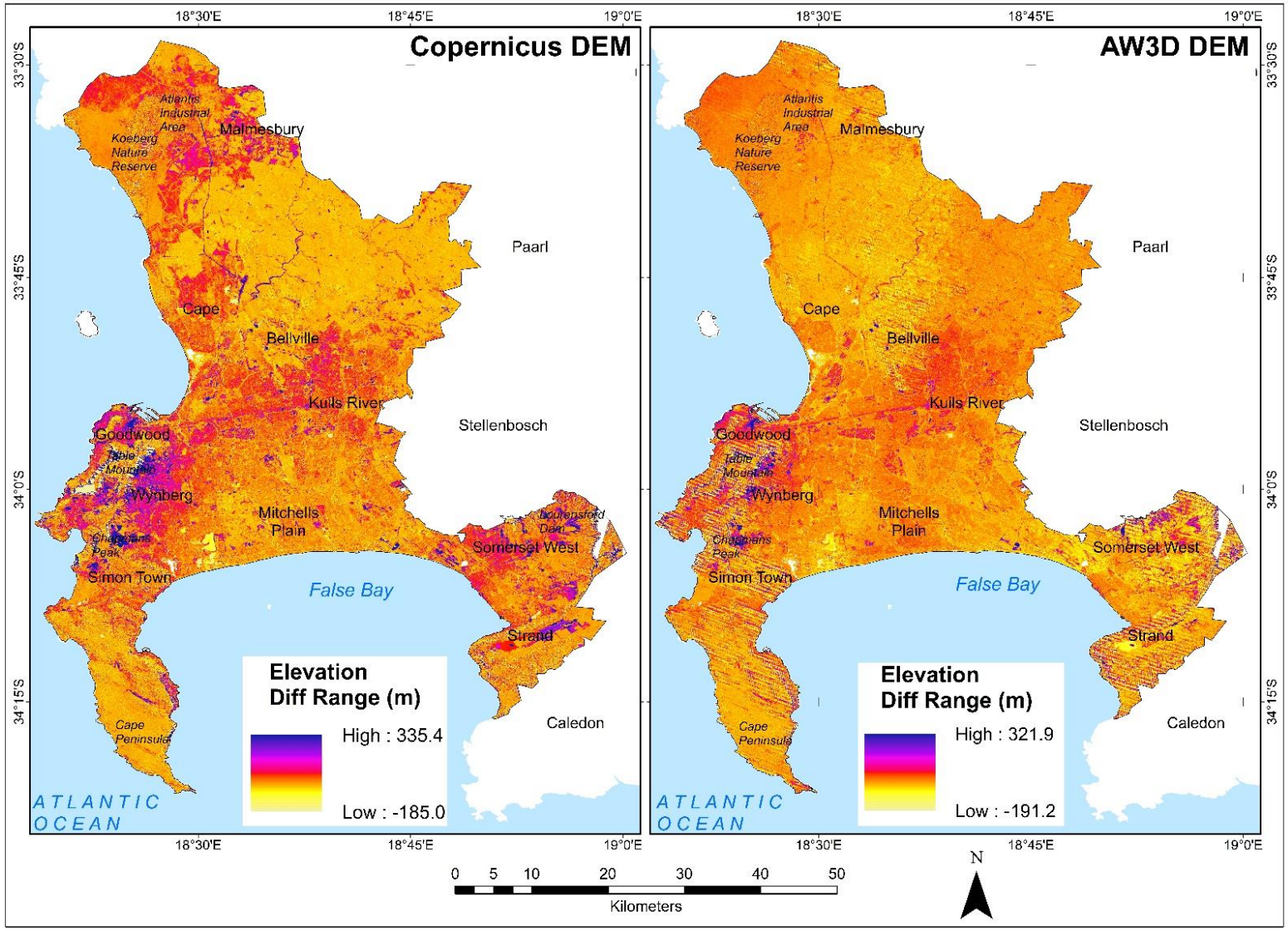


Figure 5.49: Height error maps showing the range of positive and negative elevation offsets from Copernicus DEM (left) and AW3D DEM (right)

5.4.2 Analysis of test error

Table 5.12 shows the test error comparisons of the XGBoost, LightGBM and CatBoost models trained for AW3D using default hyperparameters. Generally, the default model performed considerably well and demonstrated excellent predictive capability. Test RMSEs were lowest in the grasslands/ shrublands (1.256 – 1.273 m) and highest in the mountains (4.083 – 4.196 m). Figure 5.50 - Figure 5.52 present boxplots of the prediction error for default XGBoost (D.XGB), LightGBM (D.LGB) and CatBoost (D.Cat) respectively.

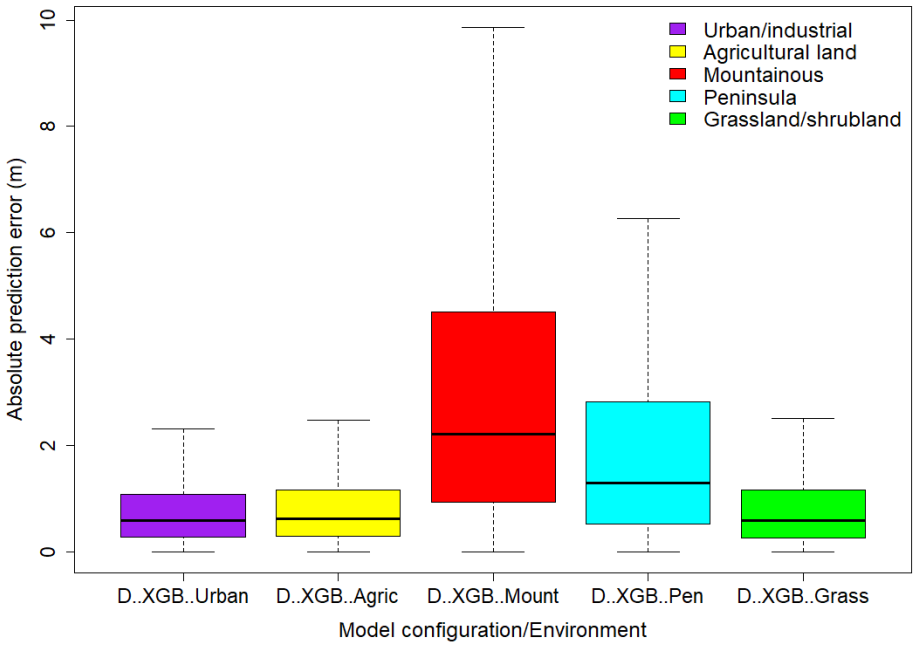


Figure 5.50: Boxplot comparison of XGBoost prediction error – AW3D

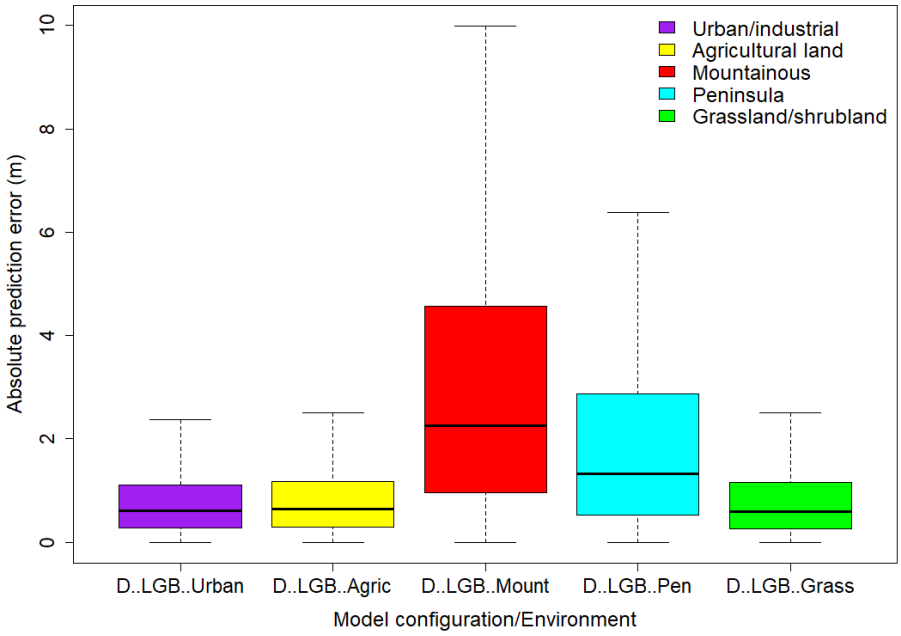


Figure 5.51: Boxplot comparison of LightGBM prediction error – AW3D

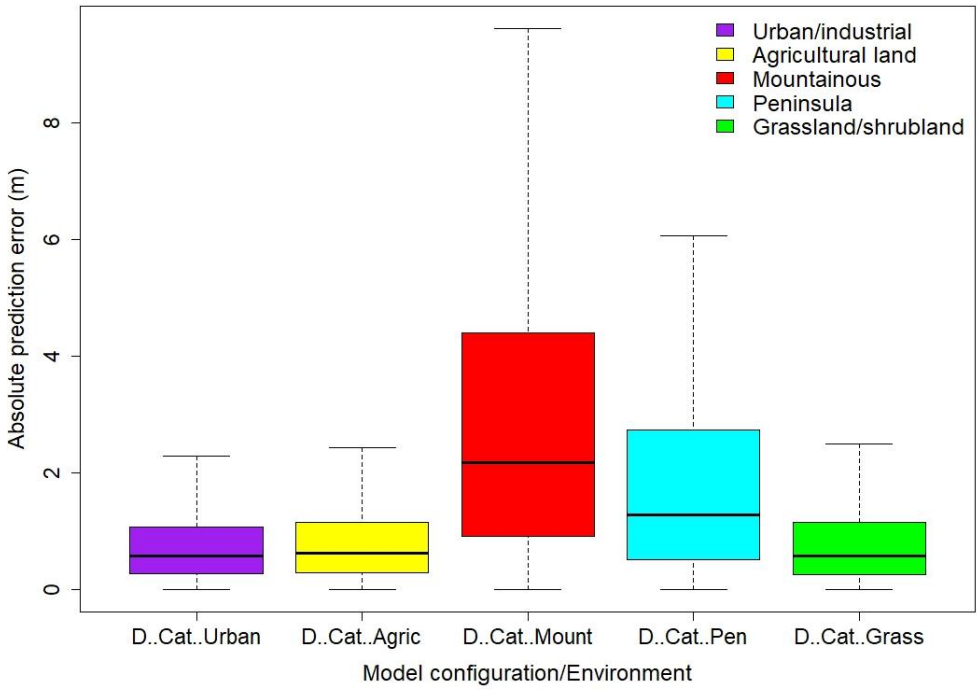


Figure 5.52: Boxplot comparison of CatBoost prediction error – AW3D

Table 5.13 - Table 5.15 show the test error comparisons of the XGBoost, LightGBM and CatBoost models trained for Copernicus using both default and optimised hyperparameters. Generally, both the default and optimised models performed considerably well and demonstrated excellent predictive capability. Test RMSEs were lowest in the agricultural lands (0.601 – 0.860 m) and highest in the mountains (1.856 – 1.934 m). Figure 5.53 - Figure 5.55 present boxplots of the prediction error for the default XGBoost (D.XGB), Bayesian-optimised XGBoost (B.XGB), default LightGBM (D.LGB), Bayesian-optimised LightGBM (B.LGB), default CatBoost (D.Cat) and Bayesian-optimised CatBoost (B.Cat), respectively.

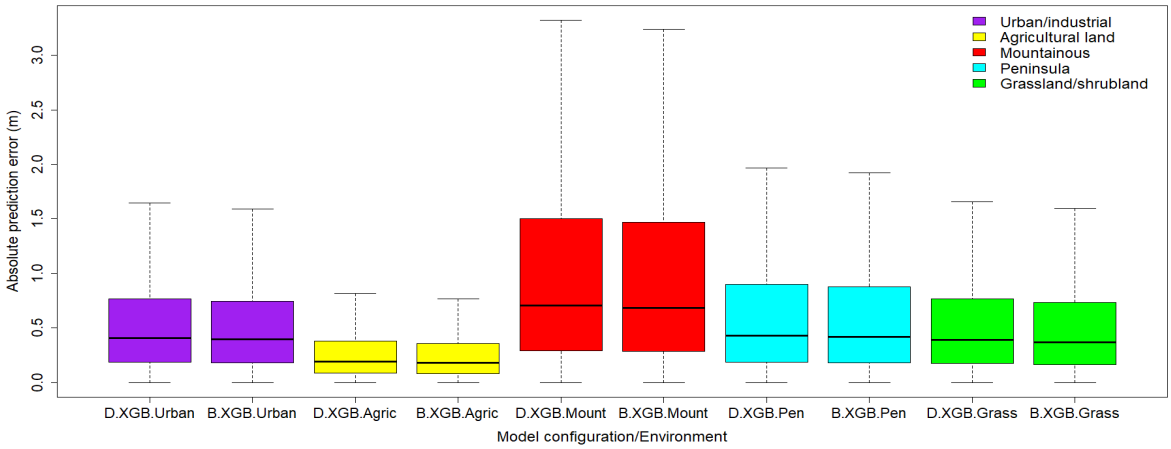


Figure 5.53: Boxplot comparison of XGBoost prediction error – Copernicus (default versus optimised hyperparameters)

Table 5.12: Comparison of XGBoost, LightGBM and CatBoost test error – AW3D (default hyperparameters). The lowest test errors are highlighted

Landscape	XGBoost			LightGBM			CatBoost		
	RMSE (m)	MAE (m)	MdAE (m)	RMSE (m)	MAE (m)	MdAE (m)	RMSE (m)	MAE (m)	MdAE (m)
Urban/ industrial	1.195	0.824	0.584	1.224	0.847	0.602	**1.179**	**0.814**	**0.577**
Agricultural land	1.311	0.889	0.621	1.330	0.903	0.635	**1.281**	**0.870**	**0.610**
Mountain	4.163	3.087	2.205	4.196	3.128	2.255	**4.083**	**3.024**	**2.170**
Peninsula	3.029	2.067	1.297	3.065	2.101	1.325	**2.940**	**2.010**	**1.272**
Grassland/ shrubland	1.274	0.871	0.581	1.273	0.874	0.586	**1.256**	**0.861**	**0.572**

Table 5.13: Comparison of XGBoost test error – Copernicus (default hyperparameters versus Bayesian-optimised hyperparameters). The lowest test errors are highlighted

Landscape	Default			Bayesian optimised		
	RMSE (m)	MAE (m)	MdAE (m)	RMSE (m)	MAE (m)	MdAE (m)
Urban/ industrial	0.893	0.597	0.409	**0.859**	**0.576**	**0.394**
Agricultural	0.648	0.342	0.191	**0.601**	**0.320**	**0.179**
Mountain	1.927	1.214	0.704	**1.883**	**1.183**	**0.685**
Peninsula	1.190	0.726	0.429	**1.138**	**0.700**	**0.419**
Grassland/ shrubland	0.864	0.584	0.392	**0.838**	**0.562**	**0.370**

Table 5.14: Comparison of LightGBM test error – Copernicus (default hyperparameters versus Bayesian-optimised hyperparameters). The lowest test errors are highlighted

Landscape	Default			Bayesian optimised		
	RMSE (m)	MAE (m)	MdAE (m)	RMSE (m)	MAE (m)	MdAE (m)
Urban/ industrial	**0.906**	**0.606**	**0.416**	0.935	0.619	0.421
Agricultural land	**0.656**	**0.346**	**0.194**	0.860	0.436	0.251
Mountain	**1.920**	**1.210**	**0.700**	1.934	1.225	0.716
Peninsula	**1.172**	**0.718**	**0.428**	1.195	0.729	0.435
Grassland/ shrubland	**0.861**	**0.585**	**0.393**	0.968	0.665	0.469

Table 5.15: Comparison of CatBoost test error – Copernicus (default hyperparameters versus Bayesian-optimised hyperparameters). The lowest test errors are highlighted

Landscape	Default			Bayesian optimised		
	RMSE (m)	MAE (m)	MdAE (m)	RMSE (m)	MAE (m)	MdAE (m)
Urban/ industrial	0.877	0.587	0.402	**0.857**	**0.573**	**0.393**
Agricultural	0.634	0.336	0.188	**0.602**	**0.319**	**0.179**
Mountain	1.908	1.203	0.695	**1.856**	**1.164**	**0.672**
Peninsula	1.170	0.717	0.428	**1.145**	**0.703**	**0.421**
Grassland/ shrubland	0.854	0.583	0.393	**0.842**	**0.568**	**0.378**

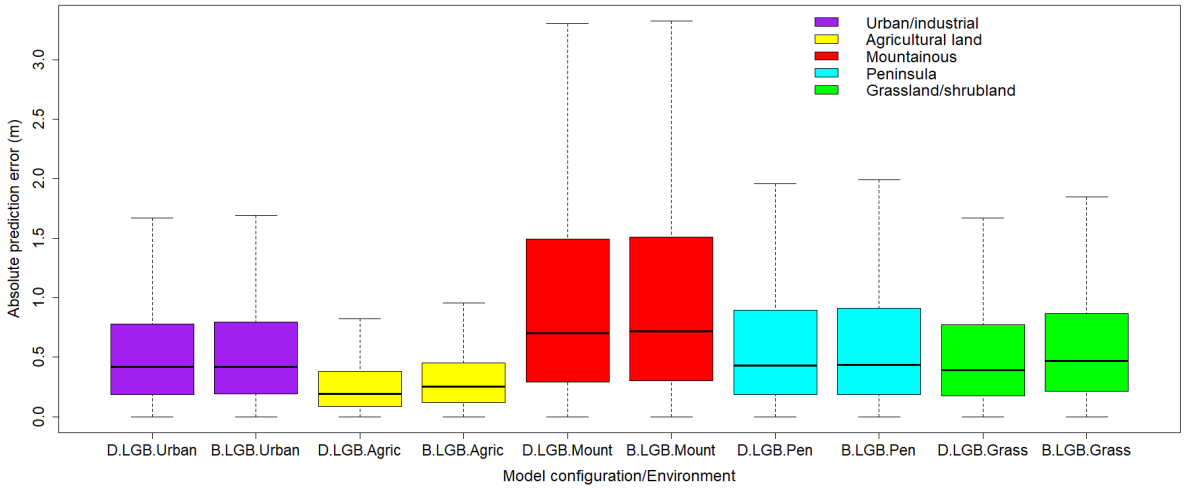


Figure 5.54: Boxplot comparison of LightGBM prediction error – Copernicus (default versus optimised hyperparameters)

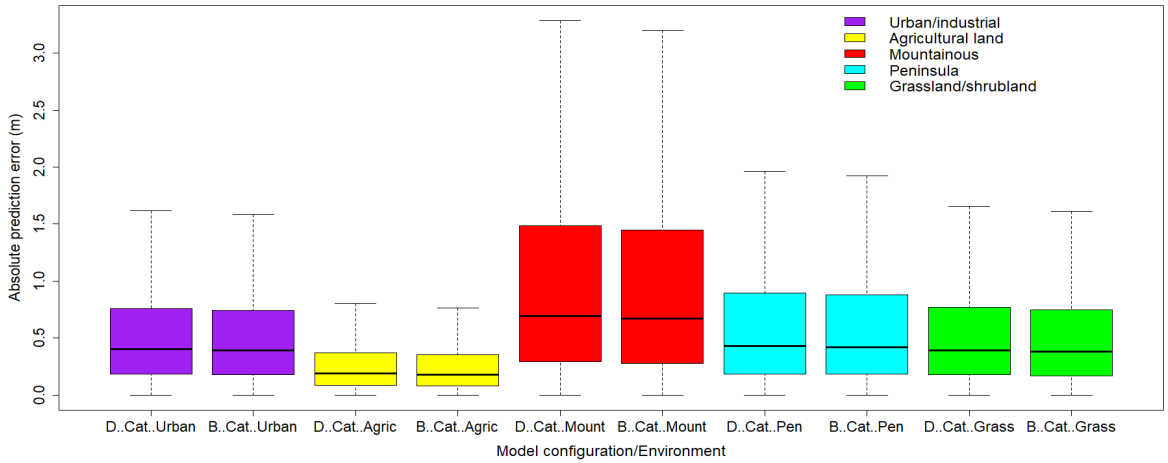


Figure 5.55: Boxplot comparison of CatBoost prediction error – Copernicus (default versus optimised hyperparameters)

5.4.3 Comparison of model execution times

Table 5.16 shows the Bayesian optimisation execution time (in seconds) of XGBoost, LightGBM and CatBoost for Copernicus DEM. LightGBM was several times faster than XGBoost and CatBoost. For example, in the Cape Peninsula, LightGBM was 15x faster than XGBoost (167.138 s vs 2486.832 s) and 25x faster than CatBoost (167.138 s vs 4102.972 s) respectively. Similarly, in the grasslands/shrublands, it was 20x faster than XGBoost (151.293 s vs 2976.456 s), and 28x faster than CatBoost (151.293 s vs 4267.270 s) respectively. The optimisation process consumed a considerable amount of time and system resources, especially in the urban/industrial areas and agricultural lands. XGBoost was slower than CatBoost in the optimisation time in the urban/industrial and agricultural lands, but was faster in the mountain, peninsula and grassland/shrubland areas. However, the number of training samples in the urban and agricultural lands exceeded that of other landscapes by several hundreds of thousands.

Table 5.16: Execution time (in seconds) for Bayesian optimisation of XGBoost, LightGBM and CatBoost – Copernicus DEM. The fastest execution times are highlighted

Landscape	Execution time for Bayesian optimisation		
	XGBoost	**LightGBM**	**CatBoost**
Urban/ industrial	17959.282	**1033.409**	6157.880
Agricultural	26080.685	**1691.126**	14862.721
Mountain	2911.410	**295.583**	4548.114
Peninsula	2486.832	**167.138**	4102.972
Grassland/ shrubland	2976.456	**151.293**	4267.270

Table 5.17 shows the execution time comparisons for model training and prediction of AW3D DEM. Generally, LightGBM had the fastest training time in all the landscapes followed by CatBoost and XGBoost. However, XGBoost surpassed CatBoost in the Cape Peninsula.

Table 5.17: Execution time (in seconds) for training and prediction of XGBoost, LightGBM and CatBoost using default hyperparameters – AW3D. The fastest execution times are highlighted

Landscape	XGBoost		LightGBM		CatBoost	
	Training	**Prediction**	**Training**	**Prediction**	**Training**	**Prediction**
Urban/ industrial	38.122	0.042	**1.892**	0.090	15.497	**0.028**
Agricultural	25.226	0.041	**1.015**	0.076	15.950	**0.030**
Mountain	29.014	**0.021**	**0.575**	0.034	8.488	**0.021**
Peninsula	7.839	**0.014**	**0.352**	0.025	22.011	0.037
Grassland/ shrubland	3.236	0.009	**0.390**	0.019	2.797	**0.007**

Table 5.18 - Table 5.20 show the execution time comparisons for model training and prediction of Copernicus DEM. The Bayesian optimisation increased the execution time of the models. For example, the XGBoost training time based on optimised hyperparameters was about 10x slower than with default hyperparameters in the urban/industrial landscape (291.885 s vs 27.956 s), and about 50x slower than with default hyperparameters in the agricultural lands. Although the prediction times were considerably less than the training time, the optimisation still slowed down the process. The training time of the optimised models was considerably slower than the default models. The optimised CatBoost was 61x slower than the default CatBoost in the urban/industrial area (916.761 s vs 14.949 s), 115x slower in the agricultural lands (1860.448 s vs 16.110 s), 50x slower in the mountainous area (1824.249 s vs 36.441 s), 142x slower in the peninsula (290.231 s vs 2.038 s) and 218x slower in the grassland/shrubland area (489.858 s vs 2.244 s). LightGBM was more efficient in most environments, with the

training time of the optimised models being only 3 – 8x slower than the default models, except for the agricultural fields.

Thus, there has to be a trade-off between the expected accuracies and the level of optimisation. However, the accuracy improvements achieved through hyperparameter optimisation could have tremendous benefits for the use of the DEMs in hydrological modelling and other environmental applications. Table 5.19 shows that even the Bayesian hyperparameter-optimised LightGBM could be trained faster than the default implementation of XGBoost and CatBoost. For example, in most cases, the training times of the optimised LightGBM model in the urban/industrial and mountain landscapes were faster than the default training times of XGBoost and CatBoost.

Table 5.18: Execution time (in seconds) for training and prediction of XGBoost using default and optimised hyperparameters – Copernicus. The fastest execution times are highlighted

Landscape	**Training**		**Prediction**	
	Default	**Bayesian optimised**	**Default**	**Bayesian optimised**
Urban/ industrial	**27.956**	291.885	**0.076**	0.399
Agricultural	**26.205**	1371.493	**0.038**	3.900
Mountain	**10.487**	106.034	**0.018**	0.221
Peninsula	**2.789**	135.030	**0.008**	0.268
Grassland/ shrubland	**3.961**	88.611	**0.017**	0.163

Table 5.19: Execution time (in seconds) for training and prediction of LightGBM using default and optimised hyperparameters – Copernicus. The fastest execution times are highlighted

Landscape	**Training**		**Prediction**	
	Default	**Bayesian optimised**	**Default**	**Bayesian optimised**
Urban/ industrial	**2.385**	20.045	**0.091**	2.603
Agricultural	**0.976**	74.576	**0.073**	5.545
Mountain	**0.597**	4.877	**0.035**	0.164
Peninsula	**0.356**	1.601	**0.022**	0.084
Grassland/ shrubland	**0.383**	1.208	**0.021**	0.108

Table 5.20: Execution time (in seconds) for training and prediction of CatBoost using default and optimised hyperparameters – Copernicus. The fastest execution times are highlighted

Landscape	Training		Prediction	
	Default	**Bayesian optimised**	**Default**	**Bayesian optimised**
Urban/ industrial	**14.949**	916.761	**0.029**	0.457
Agricultural land	**16.110**	1860.448	**0.032**	0.911
Mountain	**36.441**	1824.249	**0.016**	0.156
Peninsula	**2.038**	290.231	**0.008**	0.059
Grassland/ shrubland	**2.244**	489.858	**0.009**	0.101

5.4.4 AW3D DEM enhancement using default hyperparameters

Table 5.21 shows the accuracy measures of the corrected AW3D DEM, and a comparison with the original DEM while Table 5.22 shows the percentage reduction in MAE and RMSE after correction. Analysis of AW3D DEM corrections shows that the default models achieved reasonable performance in most landscapes and instances. The highest improvements at the 4th implementation site include: urban/industrial – MAE reduced by 78.3% from 4.462 m to 0.967 m (LightGBM) and RMSE reduced by 70.3% from 4.948 m to 1.468 m (LightGBM); agricultural land – MAE reduced by 63.5% from 2.483 m to 0.905 m (CatBoost) and RMSE reduced by 55.3% from 3.036 m to 1.358 m (CatBoost); mountain – MAE reduced by 15.5% from 7.999 m to 6.759 m (LightGBM) and RMSE reduced by 6.6% from 15.465 m to 14.440 m (CatBoost); peninsula – MAE reduced by 61.5% from 2.265 m to 0.873 m (LightGBM) and RMSE reduced by 46.6% from 2.658 m to 1.420 m (CatBoost), and the grassland/shrublands – reduction of the MAE by 81.6% from 4.085 m to 0.751 m (CatBoost) and RMSE reduced by 77.7% from 4.166 m to 0.929 m (CatBoost).

Table 5.21: Accuracy measures of the corrected AW3D DEM at the 4th implementation site, using default hyperparameters. The best accuracies are highlighted

Landscape	Original AW3D DEM		Corrected AW3D DEM (Default)					
			XGBoost		LightGBM		CatBoost	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Urban/ industrial	4.462	4.948	0.969	1.490	**0.967**	**1.468**	0.969	1.484
Agricultural land	2.483	3.036	0.910	1.361	0.910	1.374	**0.905**	**1.358**
Mountain	7.999	15.465	6.833	14.609	**6.759**	14.465	6.775	**14.440**
Peninsula	2.265	2.658	0.877	1.441	**0.873**	1.431	0.879	**1.420**
Grassland/ shrubland	4.085	4.166	0.778	0.960	0.792	0.971	**0.751**	**0.929**

Table 5.22: Percentage reduction in MAE and RMSE of AW3D DEM after correction, at the 4th implementation site, using default hyperparameters. The highest percentages are highlighted

Landscape	XGBoost		LightGBM		CatBoost	
	% MAE	% RMSE	% MAE	% RMSE	% MAE	% RMSE
Urban/ industrial	**78.3**	69.9	**78.3**	**70.3**	**78.3**	70.0
Agricultural land	63.4	55.2	63.4	54.7	**63.5**	**55.3**
Mountain	14.6	5.5	**15.5**	6.5	15.3	**6.6**
Peninsula	61.3	45.8	**61.5**	46.2	61.2	**46.6**
Grassland/ shrubland	80.9	77.0	80.6	76.7	**81.6**	**77.7**

At the 5th implementation site (Table 5.23 and Table 5.24), the highest improvements achieved in AW3D were as follows: urban/industrial – MAE reduced by 79.7% from 3.369 m to 0.684 m (CatBoost) and RMSE reduced by 72.8% from 3.636 m to 0.990 m (LightGBM); agricultural land – MAE reduced by 53.4% from 2.258 m to 1.053 m (LightGBM) and RMSE reduced by 47.4% from 2.735 m to 1.438 m (LightGBM); mountain – MAE reduced by 21% from 7.443 m to 5.877 m (CatBoost) and RMSE reduced by 17.6% from 11.442 m to 9.427 m (CatBoost); peninsula – MAE reduced by 64.1% from 2.297 m to 0.824 m (XGBoost and CatBoost) and RMSE reduced by 52.5% from 2.604 m to 1.236 m (LightGBM), and the grassland/shrublands – reduction of the MAE by 78.8% from 3.565 m to 0.754 m (LightGBM) and RMSE reduced by 73.2% from 3.660 m to 0.982 m (LightGBM).

Table 5.23: Accuracy measures of the corrected AW3D DEM at the 5th implementation site, using default hyperparameters The best accuracies are highlighted

Landscape	Original AW3D DEM		Corrected AW3D DEM (Default)					
			XGBoost		LightGBM		CatBoost	
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)
Urban/ industrial	3.369	3.636	0.692	1.009	0.685	**0.990**	**0.684**	1.001
Agricultural land	2.258	2.735	1.066	1.456	**1.053**	**1.438**	1.058	1.440
Mountain	7.443	11.442	5.902	9.453	5.913	9.504	**5.877**	**9.427**
Peninsula	2.297	2.604	**0.824**	1.244	0.825	**1.236**	**0.824**	1.251
Grassland/ shrubland	3.565	3.660	0.778	1.001	**0.754**	**0.982**	0.761	0.985

Table 5.24: Percentage reduction in MAE and RMSE of AW3D DEM after correction, at the 5th implementation site, using default hyperparameters. The highest percentages are highlighted

Landscape	XGBoost		LightGBM		CatBoost	
	% MAE	% RMSE	% MAE	% RMSE	% MAE	% RMSE
Urban/ industrial	79.5	72.2	**79.7**	**72.8**	**79.7**	72.5
Agricultural land	52.8	46.8	**53.4**	**47.4**	53.2	47.3
Mountain	20.7	17.4	20.6	16.9	**21.0**	**17.6**
Peninsula	**64.1**	52.2	**64.1**	**52.5**	**64.1**	51.9
Grassland/ shrubland	78.2	72.6	**78.8**	**73.2**	78.7	73.1

5.4.5 Copernicus DEM enhancement using default hyperparameters

Table 5.25 - Table 5.28 present the corrections achieved in Copernicus DEM using the default implementation of the models, and the highest improvements at the 4th implementation site are summarised as follows. In the urban/industrial area, the default LightGBM achieved the highest MAE reduction of 56.9% (from 1.389 m to 0.598 m) while CatBoost achieved the highest reduction in RMSE of 52.2% (from 1.952 m to 0.933 m). CatBoost also had the best performance in the agricultural, mountain and peninsula landscapes: 24% MAE reduction (from 0.456 m to 0.346 m) and 36.8% RMSE reduction (1.466 m to 0.926 m) in the agricultural lands, 5.4% MAE reduction (from 6.807 m to 6.438 m) and 0.9% RMSE reduction (from 17.339 m to 17.182 m) in the mountainous area, 15.9% MAE reduction (from 0.424 m to 0.356 m) and 22.8% RMSE reduction (from 0.905 m to 0.699 m) in the peninsula. In the grassland/shrubland, LightGBM showed the highest improvements of 66.3% MAE reduction (from 1.357 m to 0.457 m) and 62.5% RMSE reduction (from 1.520 m to 0.570 m).

Table 5.25: Accuracy measures of the corrected Copernicus DEM at the 4th implementation site, using default hyperparameters. The best accuracies are highlighted

Landscape	Original Copernicus DEM		Corrected Copernicus DEM (Default)					
			XGBoost		LightGBM		CatBoost	
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)
Urban/ industrial	1.389	1.952	0.607	0.944	**0.598**	0.935	0.604	**0.933**
Agricultural	0.456	1.466	0.352	0.932	0.347	0.940	**0.346**	**0.926**
Mountain	6.807	17.339	6.456	17.214	6.441	17.184	**6.438**	**17.182**
Peninsula	0.424	0.905	0.361	0.725	0.357	0.708	**0.356**	**0.699**
Grassland/ shrubland	1.357	1.520	0.472	0.592	**0.457**	**0.570**	0.459	0.571

Table 5.26: Percentage reduction in MAE and RMSE of Copernicus DEM after correction, at the 4th implementation site, using default hyperparameters. The highest percentages are highlighted

Landscape	XGBoost		LightGBM		CatBoost	
	% MAE	% RMSE	% MAE	% RMSE	% MAE	% RMSE
Urban/ industrial	56.3	51.6	**56.9**	52.1	56.5	**52.2**
Agricultural	22.9	36.4	23.8	35.9	**24.0**	**36.8**
Mountain	5.2	0.7	**5.4**	**0.9**	**5.4**	**0.9**
Peninsula	14.7	19.9	15.8	21.8	**15.9**	**22.8**
Grassland/ shrubland	65.2	61.0	**66.3**	**62.5**	66.2	62.4

At the 5th implementation site (Table 5.27, Table 5.28), the default hyperparameters also achieved significant improvements in the accuracy of Copernicus DEM. The highest improvements achieved include: urban/industrial area – there was a 52.9% reduction in MAE from 1.087 m to 0.512 m (CatBoost) and a 47% reduction in RMSE from 1.479 m to 0.783 m (CatBoost); agricultural lands – 23.9% reduction in MAE from 0.420 m to 0.320 m (CatBoost) and 37.9% reduction in RMSE from 1.230 m to 0.764 m (CatBoost); mountainous area – 16.2% reduction in MAE from 4.848 m to 4.064 m (CatBoost) and 10.3% reduction in RMSE from 10.237 m to 9.187 m (CatBoost); peninsula: 14.4% reduction in MAE from 0.345 m to 0.295 m (CatBoost) and 16.1% reduction in RMSE from 0.521 m to 0.437 m (CatBoost); grassland/shrubland: 48.1% reduction in MAE from 0.497 m to 0.258 m (LightGBM) and 35.3% reduction in RMSE from 0.563 m to 0.364 m (LightGBM).

Table 5.27: Accuracy measures of the corrected Copernicus DEM at the 5th implementation site, using default hyperparameters. The best accuracies are highlighted

Landscape	Original Copernicus DEM		Corrected Copernicus DEM (Default)					
			XGBoost		LightGBM		CatBoost	
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)
Urban/ industrial	1.087	1.479	0.518	0.787	0.519	0.789	**0.512**	**0.783**
Agricultural	0.420	1.230	0.324	0.770	0.325	0.776	**0.320**	**0.764**
Mountain	4.848	10.237	4.092	9.213	4.084	9.233	**4.064**	**9.187**
Peninsula	0.345	0.521	0.303	0.448	0.301	0.445	**0.295**	**0.437**
Grassland/ shrubland	0.497	0.563	0.267	0.378	**0.258**	**0.364**	0.276	0.381

Table 5.28: Percentage reduction in MAE and RMSE of Copernicus DEM after correction, at the 5th implementation site, using default hyperparameters. The highest percentages are highlighted

Landscape	**XGBoost**		**LightGBM**		**CatBoost**	
	% MAE	**% RMSE**	**% MAE**	**% RMSE**	**% MAE**	**% RMSE**
Urban/ industrial	52.4	46.8	52.3	46.7	**52.9**	**47.0**
Agricultural	23.0	37.4	22.8	36.9	**23.9**	**37.9**
Mountain	15.6	10.0	15.8	9.8	**16.2**	**10.3**
Peninsula	12.0	14.1	12.6	14.6	**14.4**	**16.1**
Grassland/ shrubland	46.3	32.9	**48.1**	**35.3**	44.4	32.3

Figure 5.56 - Figure 5.65 showcase comparisons of height error maps and elevation surfaces of the original and corrected Copernicus DEMs. In most cases, the confounding anomalies caused by terrain offsets in the original DEMs were de-escalated in the corrected DEMs. While most of the errors were neutralised, it appears that surficial patterns or alignments could act as barriers to effective error correction as observed on the eastern flank of the urban/industrial area in Figure 5.56. In the agricultural lands (Figure 5.58), numerous terrain offsets degraded the accuracy of the original DEM. Such offsets could act as artificial blockages to surface flow in hydraulic simulation and modelling. However, the correction achieved extensive diminution of these offsets, including the offsets occurring along the Diep River floodplain.

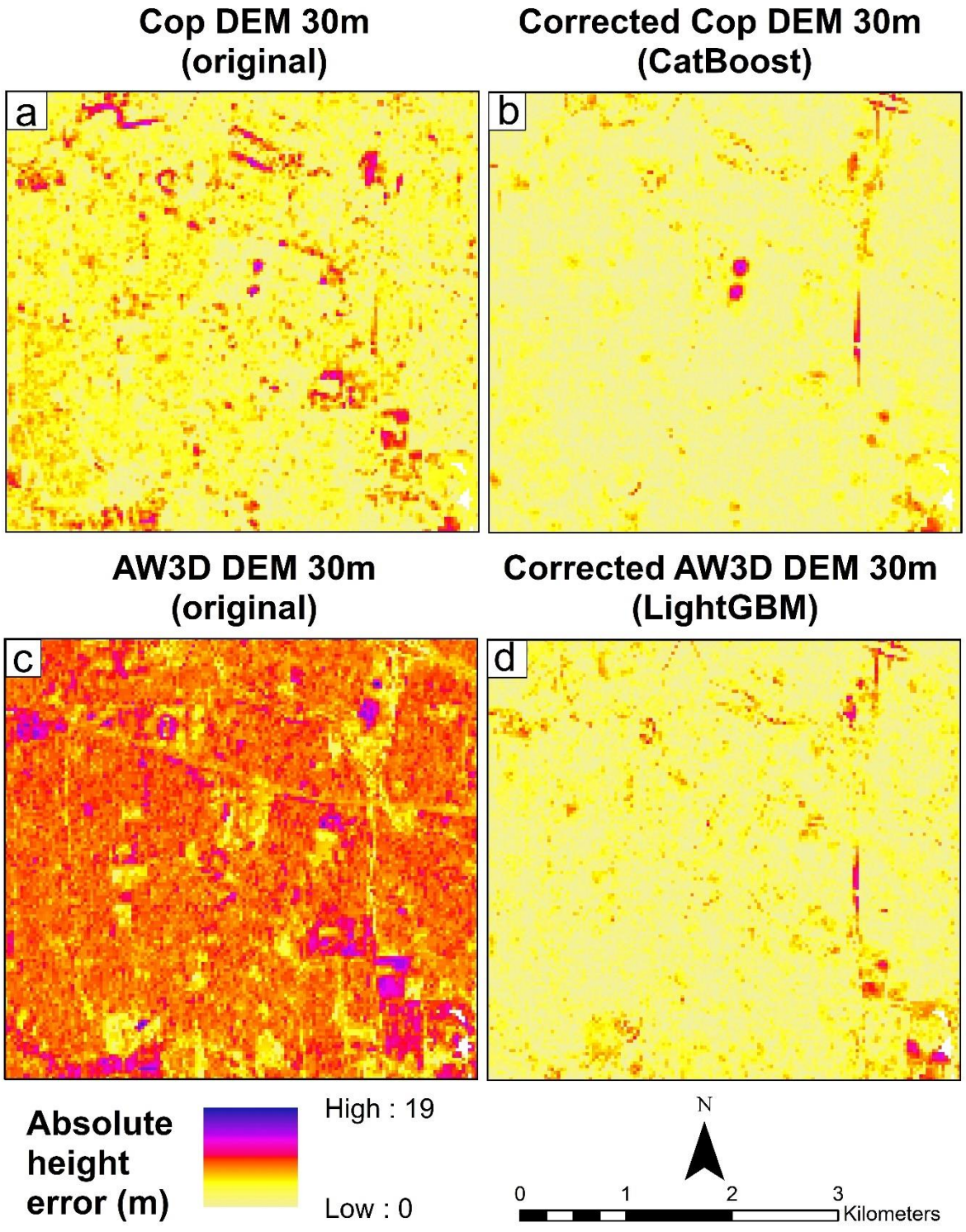


Figure 5.56: Absolute height errors of the best enhanced DEMs in the urban/industrial area, using default hyperparameters (5th implementation site) – (a) original Copernicus DEM (b) corrected Copernicus DEM (c) original AW3D DEM (d) corrected AW3D DEM.

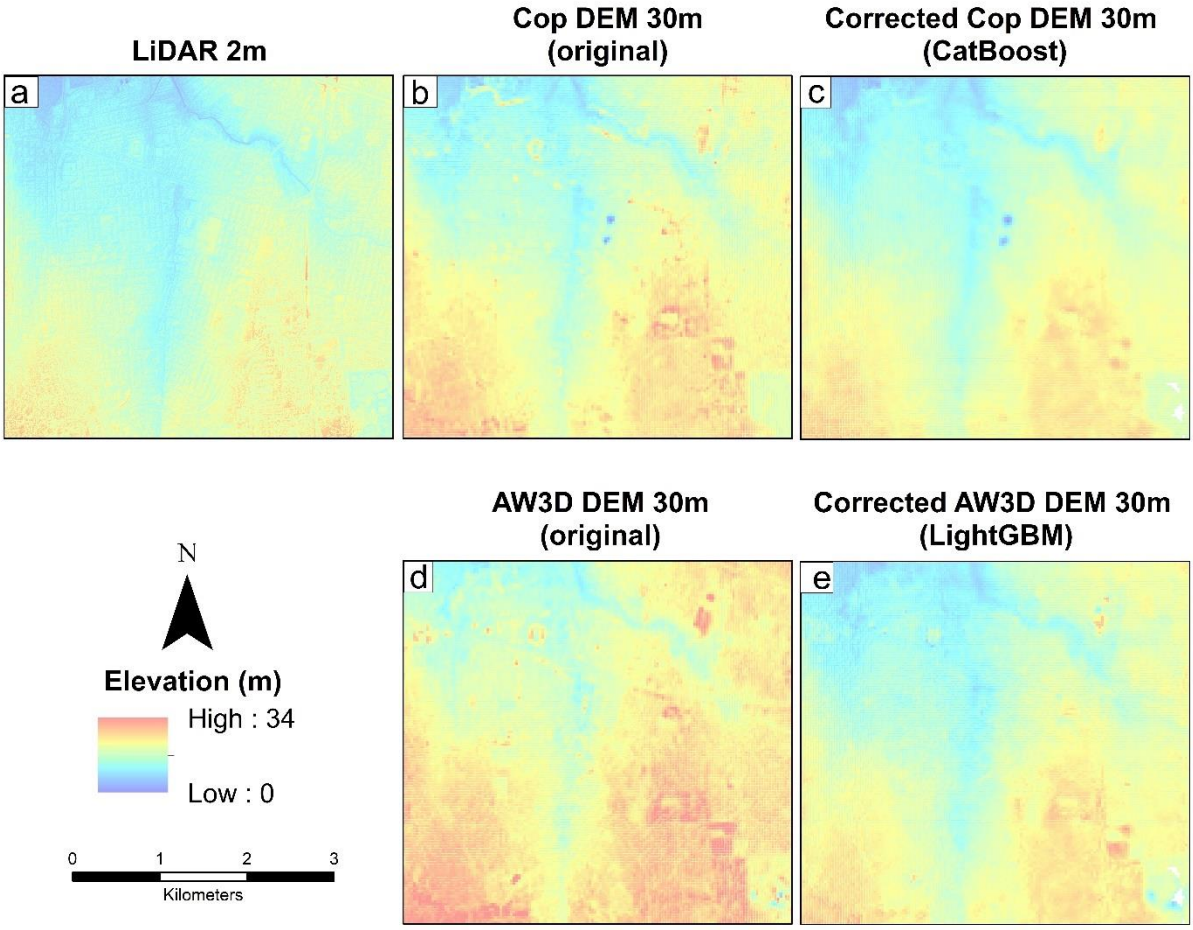


Figure 5.57: Visualisation of the best enhanced DEMs in the urban/industrial area, using default hyperparameters (5th implementation site) – (a) LiDAR DEM (b) original Copernicus DEM (c) corrected Copernicus DEM (d) original AW3D DEM (e) corrected AW3D DEM

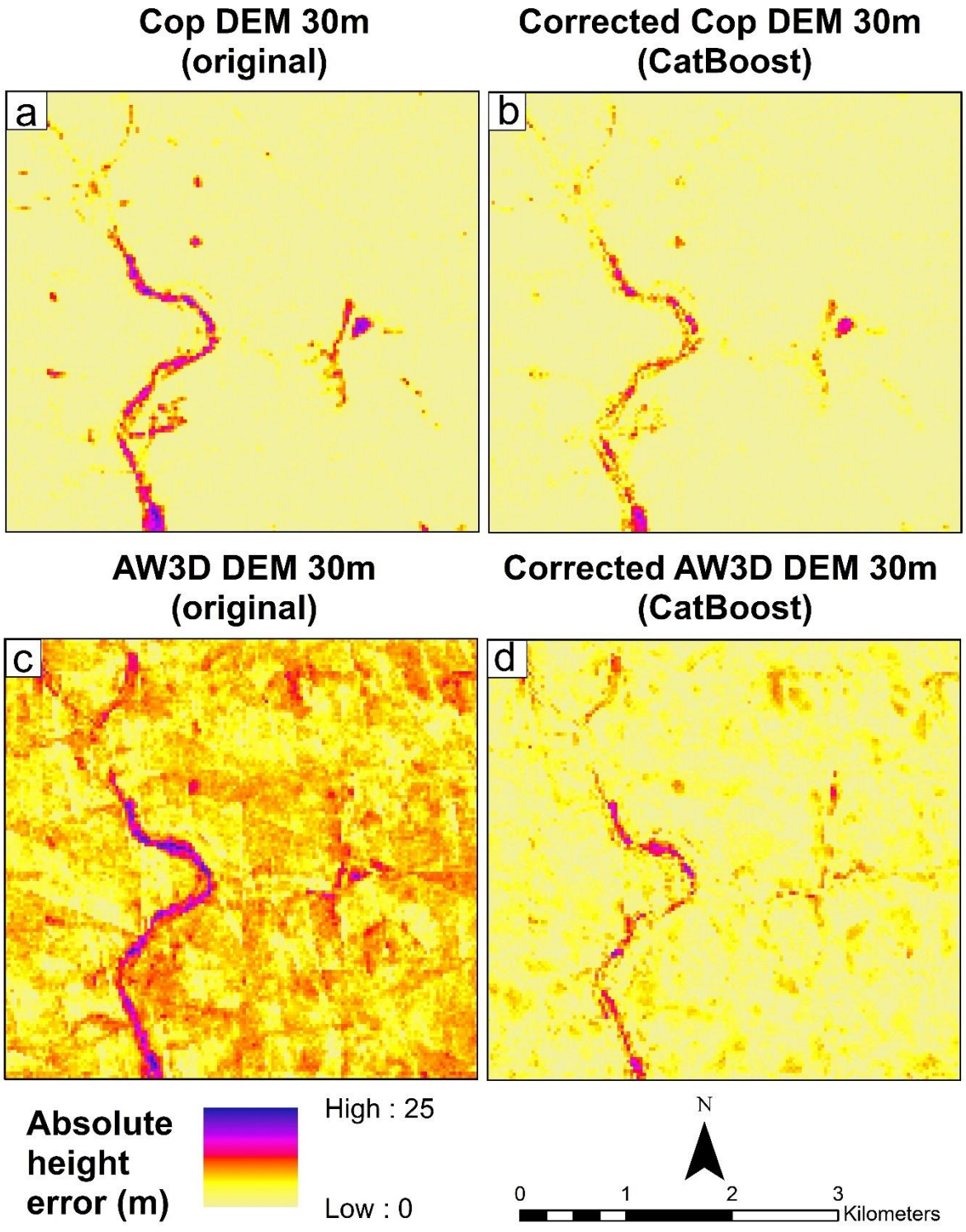


Figure 5.58: Absolute height errors of the best enhanced DEMs in the agricultural lands, using default hyperparameters (4th implementation site) – (a) original Copernicus DEM (b) corrected Copernicus DEM (c) original AW3D DEM (d) corrected AW3D DEM

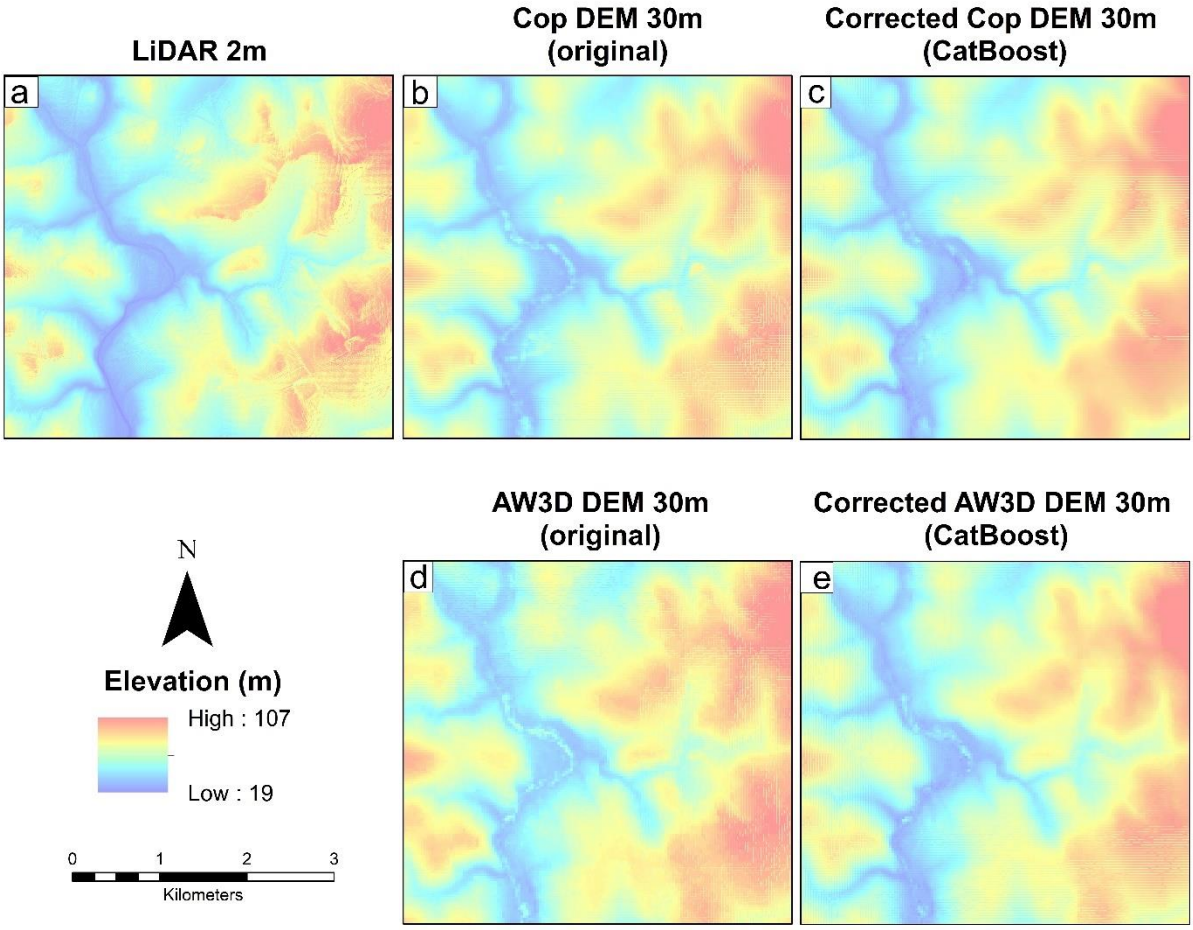


Figure 5.59: Visualisation of the best enhanced DEMs in the agricultural lands, using default hyperparameters (4th implementation site) – (a) LiDAR DEM (b) original Copernicus DEM (c) corrected Copernicus DEM (d) original AW3D DEM (e) corrected AW3D DEM

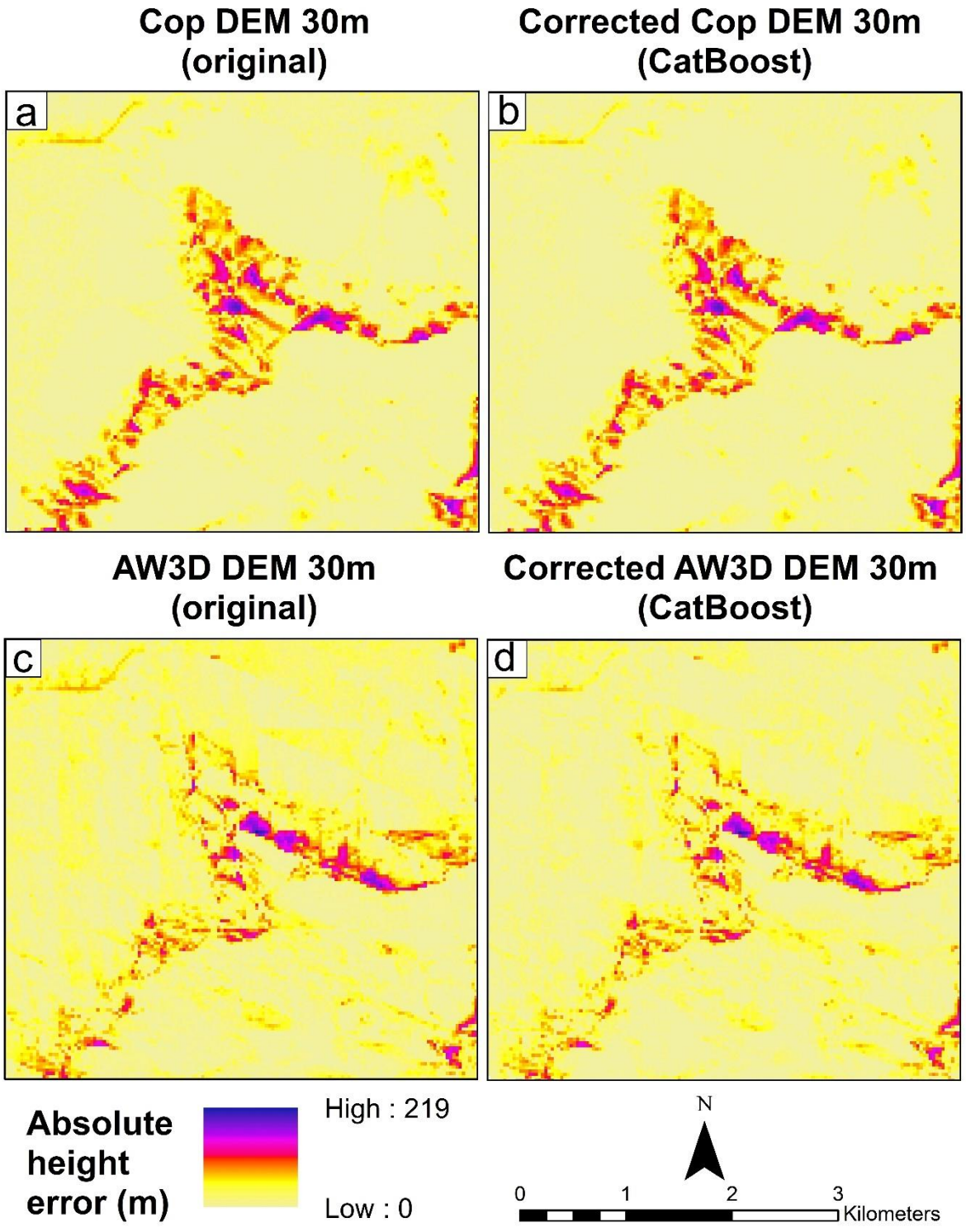


Figure 5.60: Absolute height errors of the best enhanced DEMs in the mountainous area, using default hyperparameters (4th implementation site) – (a) original Copernicus DEM (b) corrected Copernicus DEM (c) original AW3D DEM (d) corrected AW3D DEM

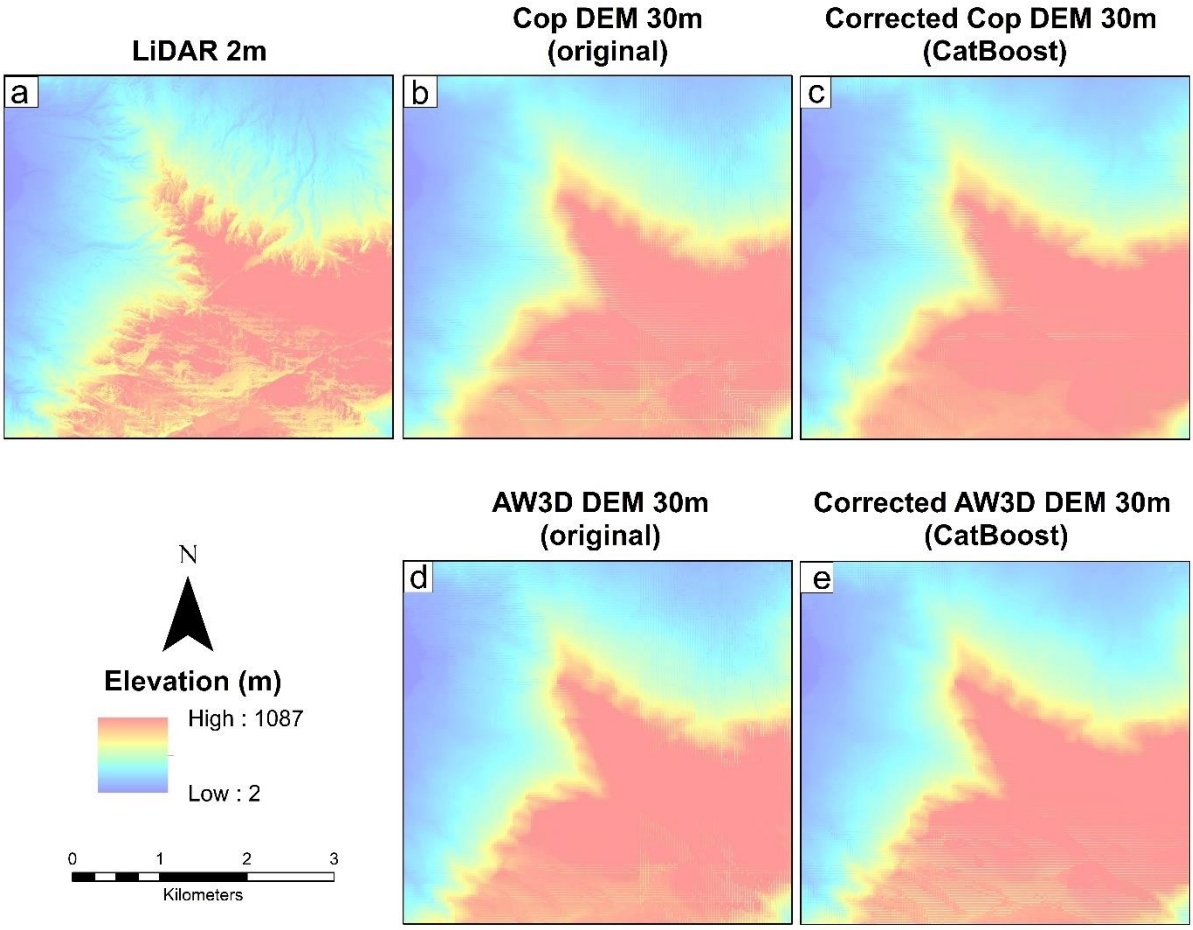


Figure 5.61: Visualisation of the best enhanced DEMs in the mountainous area, using default hyperparameters (4th implementation site) – (a) LiDAR DEM (b) original Copernicus DEM (c) corrected Copernicus DEM (d) original AW3D DEM (e) corrected AW3D DEM

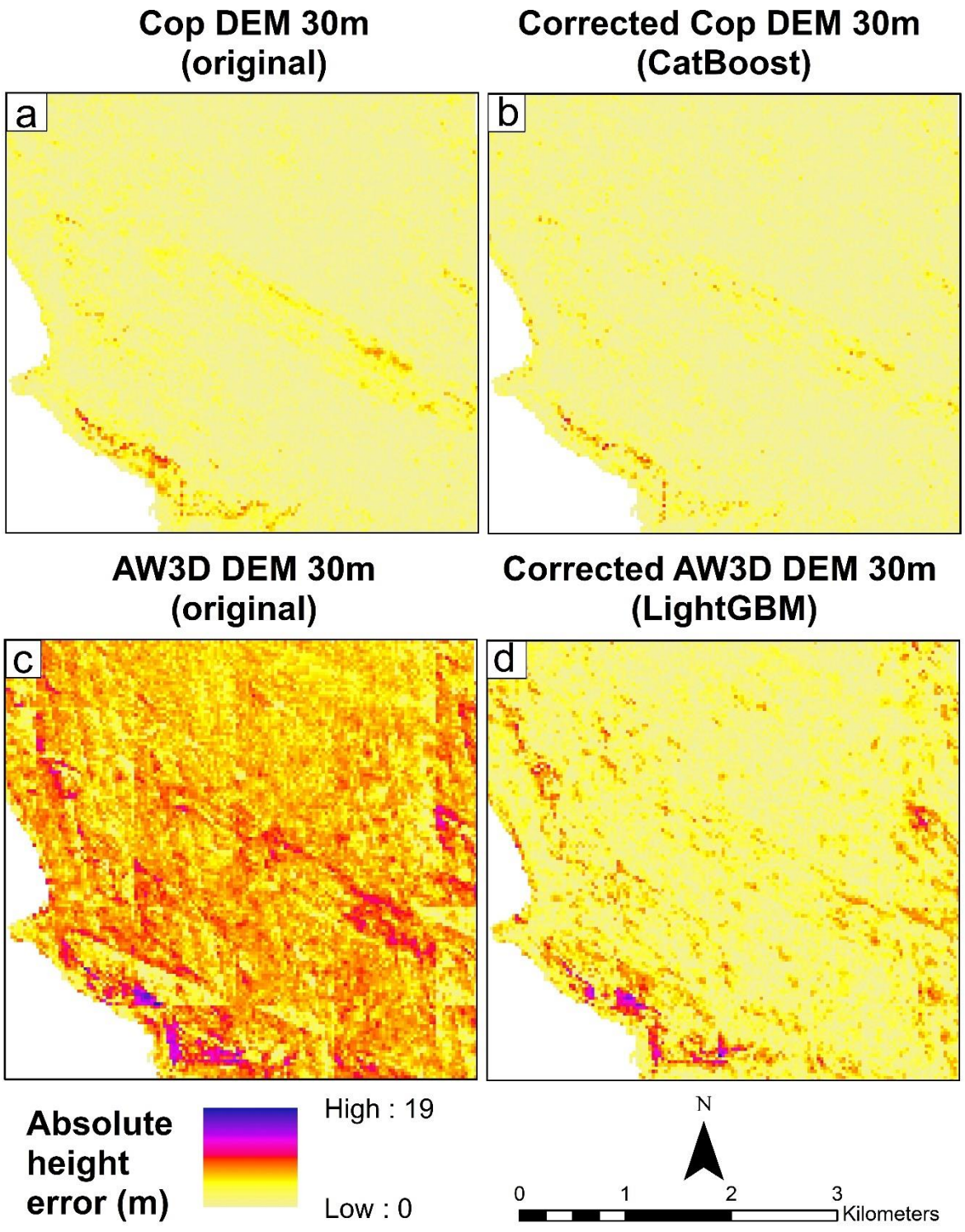


Figure 5.62: Absolute height errors of the best enhanced DEMs in the peninsula, using default hyperparameters (5th implementation site) – (a) original Copernicus DEM (b) corrected Copernicus DEM (c) original AW3D DEM (d) corrected AW3D DEM

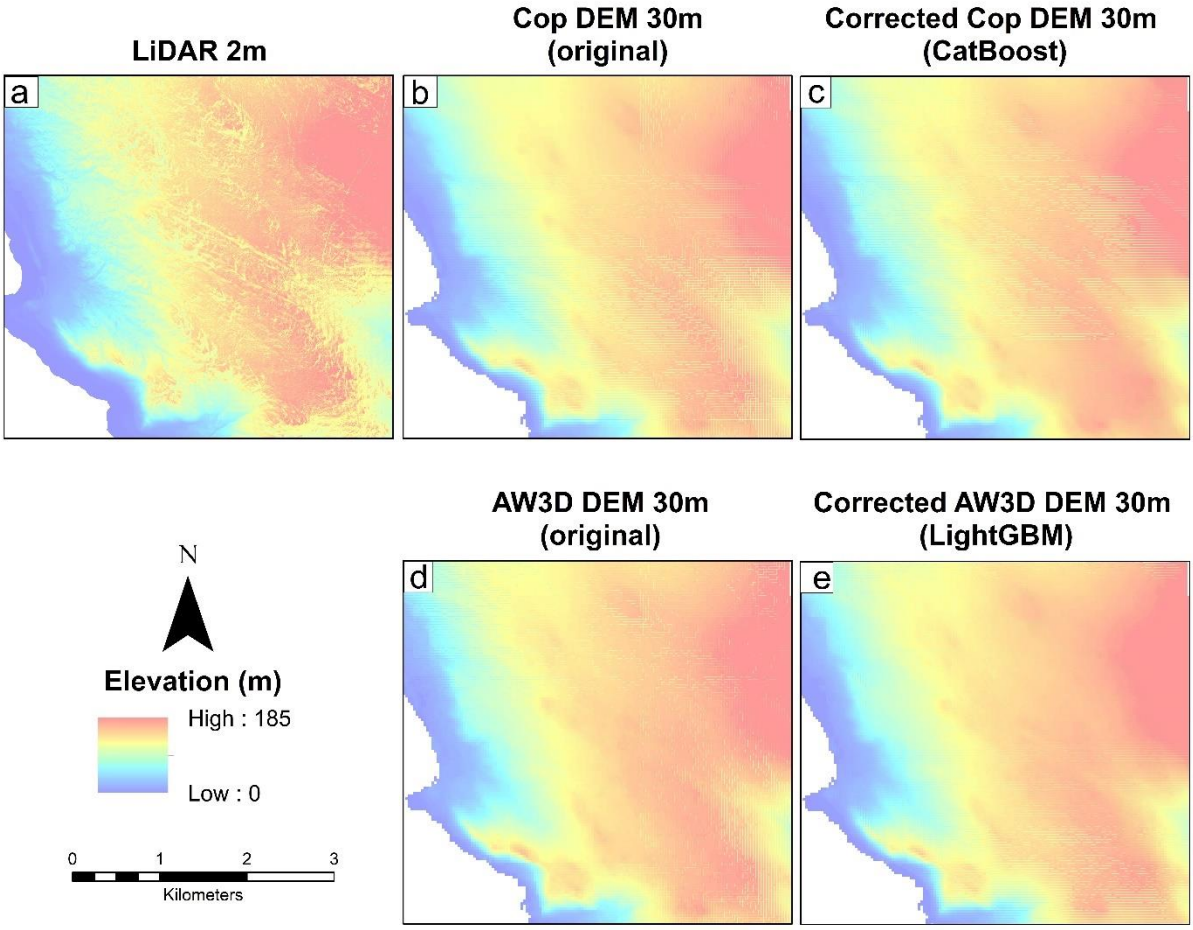


Figure 5.63: Visualisation of the best enhanced DEMs in the peninsula, using default hyperparameters (5th implementation site) – (a) LiDAR DEM (b) original Copernicus DEM (c) corrected Copernicus DEM (d) original AW3D DEM (e) corrected AW3D DEM

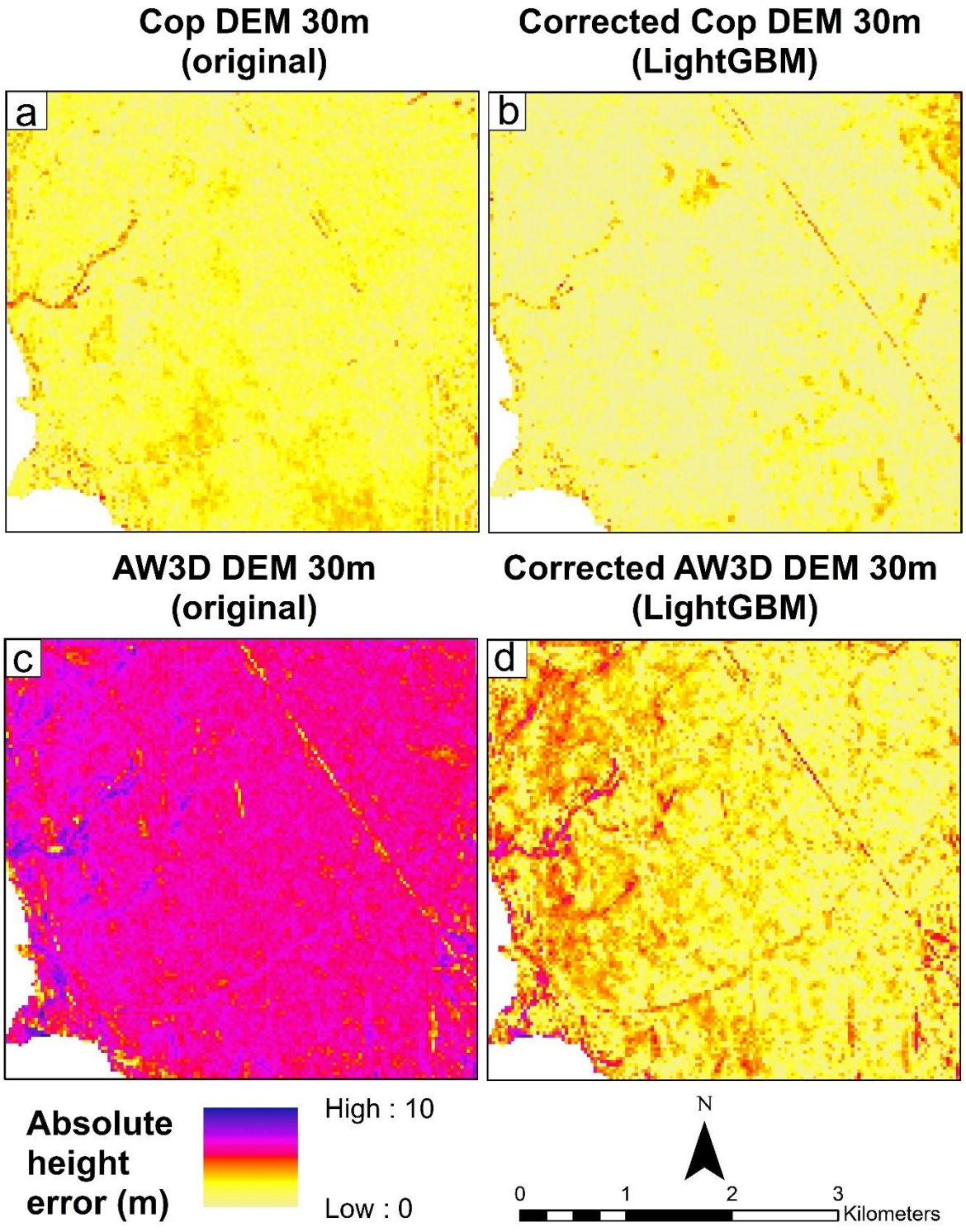


Figure 5.64: Absolute height errors of the best enhanced DEMs in the grassland/shrubland, using default hyperparameters (5th implementation site) – (a) original Copernicus DEM (b) corrected Copernicus DEM (c) original AW3D DEM (d) corrected AW3D DEM

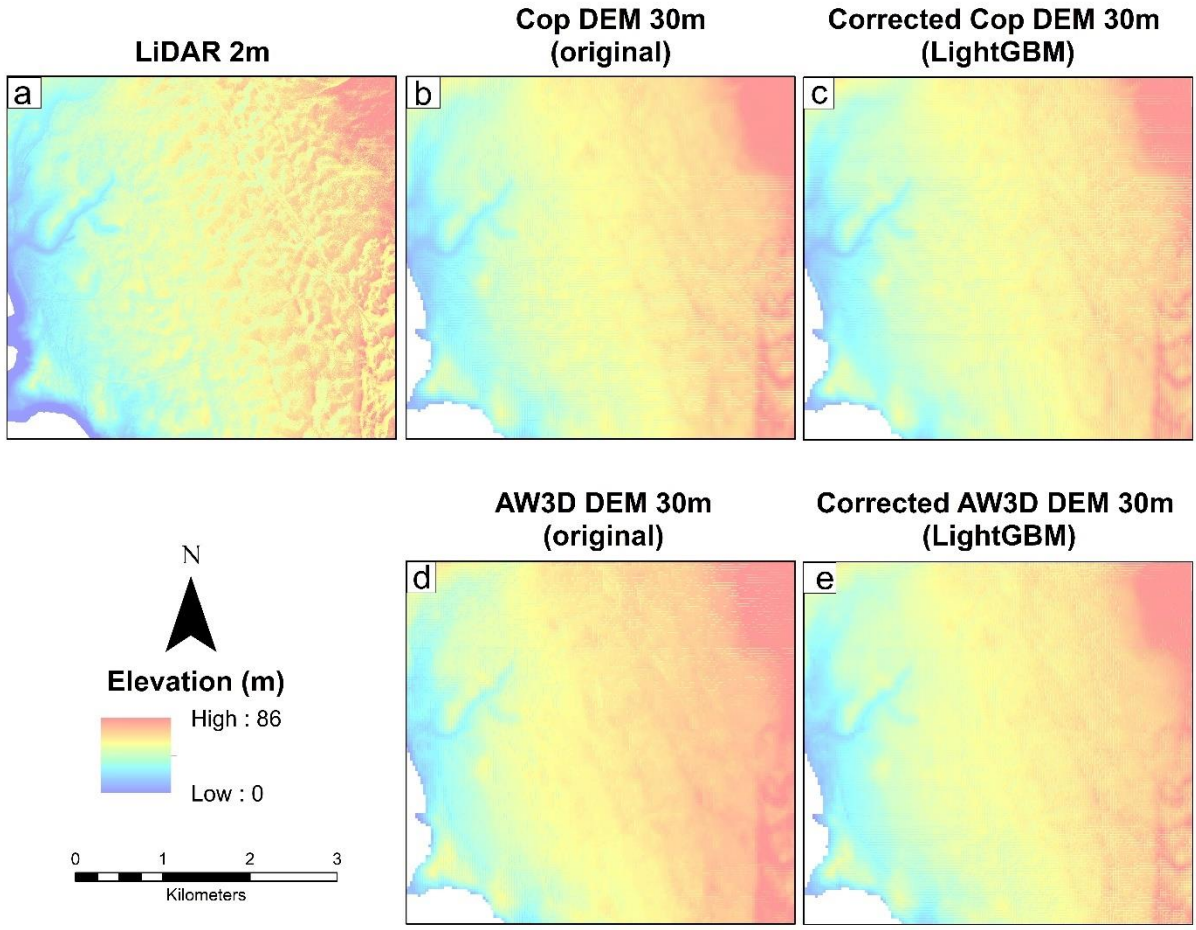


Figure 5.65: Visualisation of the best enhanced DEMs in the grassland/shrubland, using default hyperparameters (5th implementation site) – (a) LiDAR DEM (b) original Copernicus DEM (c) corrected Copernicus DEM (d) original AW3D DEM (e) corrected AW3D DEM

5.4.6 Bayesian-optimised enhancement of Copernicus DEM

This section presents the results of Bayesian optimisation of XGBoost, LightGBM and CatBoost for the enhancement of Copernicus DEM. This analysis compares the outcomes of optimisation and demonstrates its value over the default model implementations. The results from the 4th implementation site (Table 5.29, Table 5.30) show that the improvements achieved with the optimised model are highly competitive, i.e., urban/industrial - MAE reduced by 57.1% from 1.389 m to 0.596 m (LightGBM) and RMSE reduced by 51.8% from 1.952 m to 0.941 m (CatBoost); agricultural lands - MAE reduced by 24.1% from 0.456 m to 0.346 m (CatBoost) and RMSE reduced by 38.9% from 1.466 m to 0.895 m (CatBoost); mountainous area - MAE reduced by 5.5% from 6.807 m to 6.434 m (LightGBM) and RMSE reduced by 0.9% from 17.339 m to 17.179 m (XGBoost); peninsula - MAE reduced by 16.9% from 0.424 m to 0.352 m (CatBoost) and RMSE reduced by 23.4% from 0.905 m to 0.693 m (CatBoost); and grassland/shrubland - MAE reduced by 68% from 1.357 m to 0.434 m (CatBoost) and RMSE reduced by 63.8% from 1.520 m to 0.550 m (CatBoost).

Table 5.29: Accuracy measures of the corrected Copernicus DEM at the 4th implementation site, using Bayesian-optimised hyperparameters. The best accuracies are highlighted

Landscape	**Original Copernicus DEM**		**Corrected Copernicus DEM (Bayesian optimised)**					
			XGBoost		**LightGBM**		**CatBoost**	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Urban/ industrial	1.389	1.952	0.615	0.955	**0.596**	0.948	0.608	**0.941**
Agricultural land	0.456	1.466	0.355	0.937	0.463	1.330	**0.346**	**0.895**
Mountain	6.807	17.339	6.445	**17.179**	**6.434**	17.187	6.438	17.193
Peninsula	0.424	0.905	0.358	0.716	0.363	0.732	**0.352**	**0.693**
Grassland/ shrubland	1.357	1.520	0.475	0.603	0.544	0.663	**0.434**	**0.550**

Figure 5.66 - Figure 5.70 show comparisons of the absolute height error of the original Copernicus DEM versus the corrected Copernicus DEM for different landscapes at the 4th implementation site.

Table 5.30: Percentage reduction in MAE and RMSE of Copernicus DEM after correction, at the 4th implementation site, using Bayesian-optimised hyperparameters. The highest percentages are highlighted

	XGBoost		**LightGBM**		**CatBoost**	
Landscape	**% MAE**	**% RMSE**	**% MAE**	**% RMSE**	**% MAE**	**% RMSE**
Urban/industrial	55.7	51.1	**57.1**	51.5	56.2	**51.8**
Agricultural	22.2	36.1	-1.6	9.3	**24.1**	**38.9**
Mountain	5.3	**0.9**	**5.5**	**0.9**	5.4	0.8
Peninsula	15.5	20.8	14.3	19.1	**16.9**	**23.4**
Grassland/ shrubland	65.0	60.3	59.9	56.4	**68.0**	**63.8**

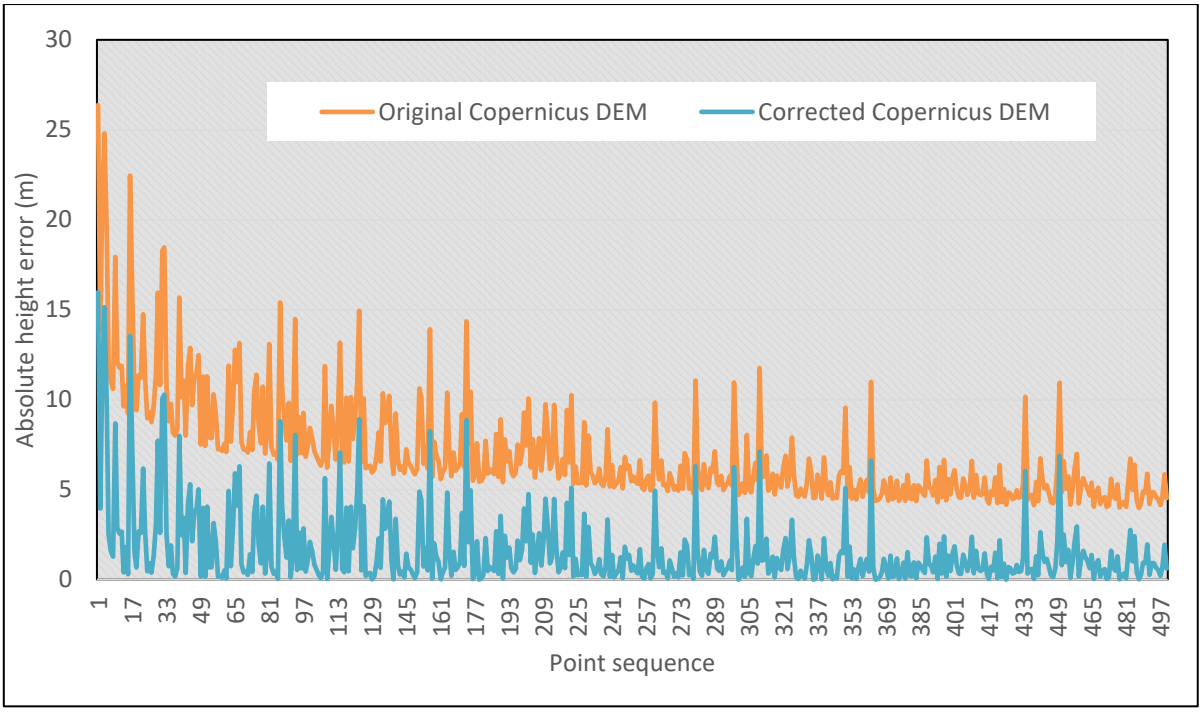


Figure 5.66: Absolute height errors at the 4th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the urban/ industrial landscape

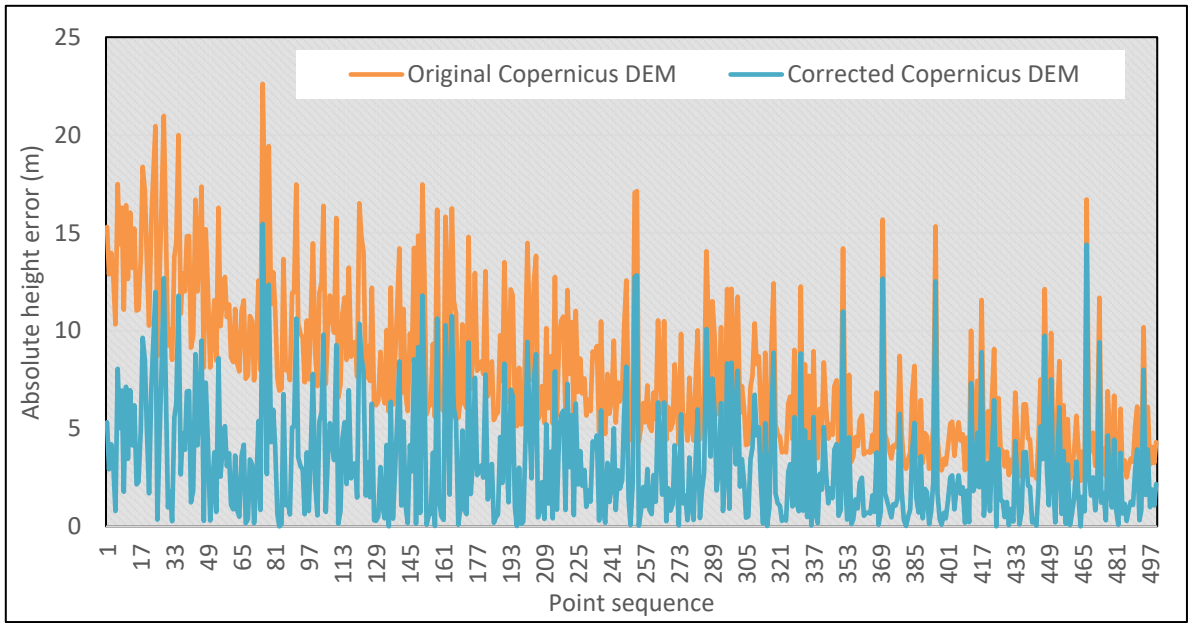


Figure 5.67: Absolute height errors at the 4th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the agricultural landscape

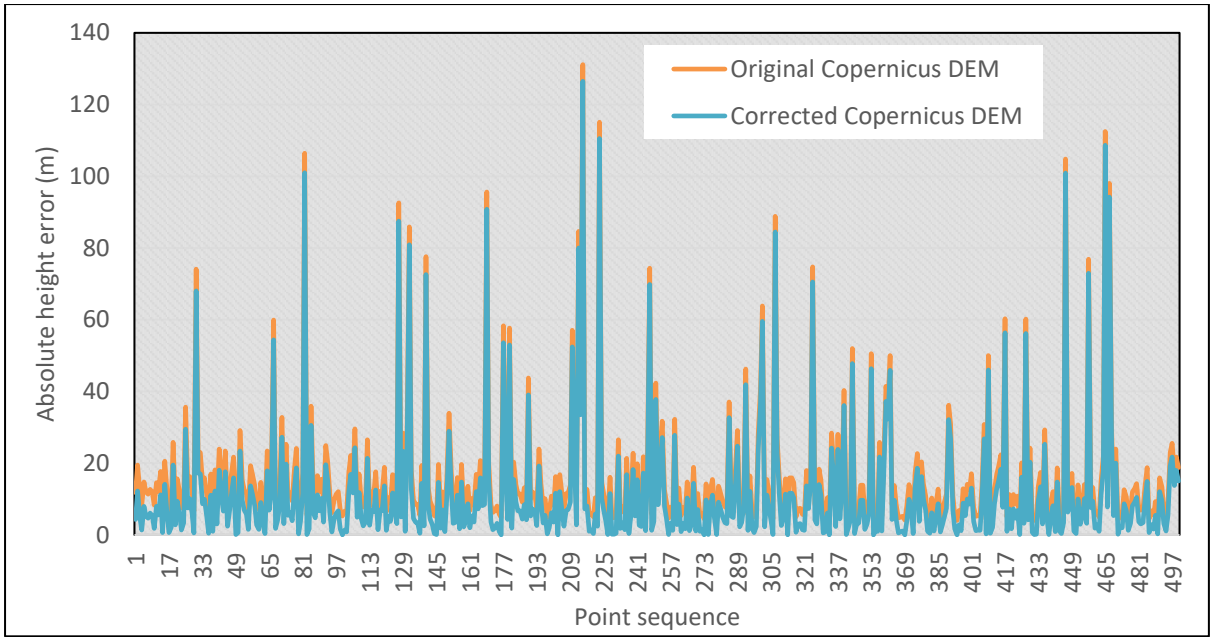


Figure 5.68: Absolute height errors at the 4th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.XGB) for 500 selected points in the mountain landscape

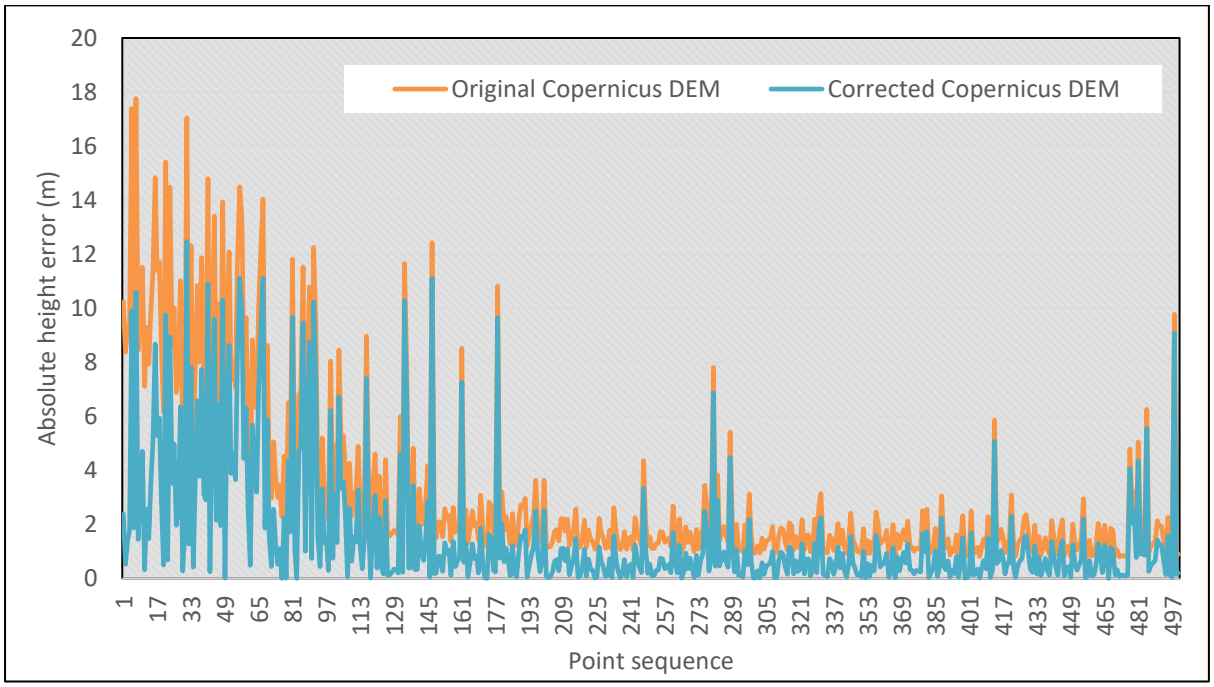


Figure 5.69: Absolute height errors at the 4th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the peninsula landscape

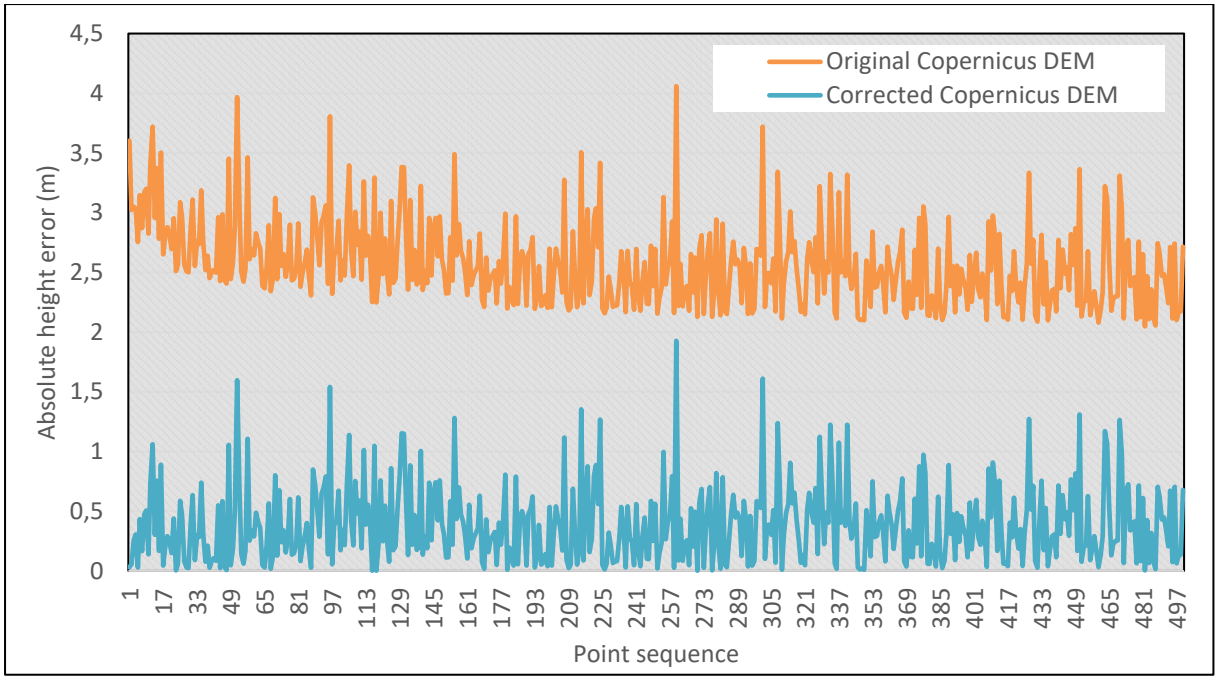


Figure 5.70: Absolute height errors at the 4th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the grassland/shrubland landscape

At the 5th implementation site, the optimised models also outperformed the default models in most landscapes and instances (Table 5.31 and Table 5.32), i.e., urban/industrial - MAE reduced by 53.1% from 1.087 m to 0.509 m (XGBoost and CatBoost) and RMSE reduced by 47.2% from 1.479 m to 0.781 m (XGBoost); peninsula - MAE reduced by 15.3% from 0.345 m to 0.292 m (CatBoost) and RMSE reduced by 17.4% from 0.521 m to 0.430 m (CatBoost); and grassland/shrubland - MAE reduced by 51.3% from 0.497 m to 0.242 m (XGBoost) and RMSE reduced by 37.4% from 0.563 m to 0.353 m (XGBoost). However, the default CatBoost model outperformed the optimised version, in the mountain landscape, achieving 16.2% MAE (from 4.848 m to 4.064 m) and 10.3% RMSE (from 10.237 m to 9.187 m) reductions compared with 15.8% MAE (from 4.848 m to 4.080 m) and 10.1% RMSE (from 10.237 m to 9.206 m) reductions by the optimised model. In the agricultural landscape, the highest reduction of 23.9% in MAE from 0.420 m to 0.320 m was achieved by the default CatBoost model while the highest reduction in RMSE of 38.2% from 1.230 m to 0.760 m (CatBoost) was achieved through optimisation.

Figure 5.71 - Figure 5.75 show comparisons of the absolute height error of the original Copernicus DEM versus the corrected Copernicus DEM for different landscapes at the 5th implementation site.

Table 5.31: Accuracy measures of the corrected Copernicus DEM at the 5th implementation site, using Bayesian-optimised hyperparameters. The best accuracies are highlighted

Landscape	Original Copernicus DEM		Corrected Copernicus DEM (Bayesian optimised)					
			XGBoost		LightGBM		CatBoost	
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)
Urban/ industrial	1.087	1.479	**0.509**	**0.781**	0.517	0.785	**0.509**	0.788
Agricultural	0.420	1.230	0.326	0.784	0.434	1.122	**0.325**	**0.760**
Mountain	4.848	10.237	4.127	9.270	4.141	9.321	**4.080**	**9.206**
Peninsula	0.345	0.521	0.297	0.437	0.308	0.459	**0.292**	**0.430**
Grassland/ shrubland	0.497	0.563	**0.242**	**0.353**	0.313	0.400	0.266	0.373

Table 5.32: Percentage reduction in MAE and RMSE of Copernicus DEM after correction, at the 5th implementation site, using Bayesian-optimised hyperparameters. The highest percentages are highlighted

Landscape	XGBoost		LightGBM		CatBoost	
	% MAE	% RMSE	% MAE	% RMSE	% MAE	% RMSE
Urban/ industrial	**53.1**	**47.2**	52.4	46.9	**53.1**	46.7
Agricultural land	22.4	36.2	-3.3	8.7	**22.8**	**38.2**
Mountain	14.9	9.4	14.6	9.0	**15.8**	**10.1**
Peninsula	13.7	16.1	10.5	12.0	**15.3**	**17.4**
Grassland/ shrubland	**51.3**	**37.4**	37.1	28.9	46.6	33.8

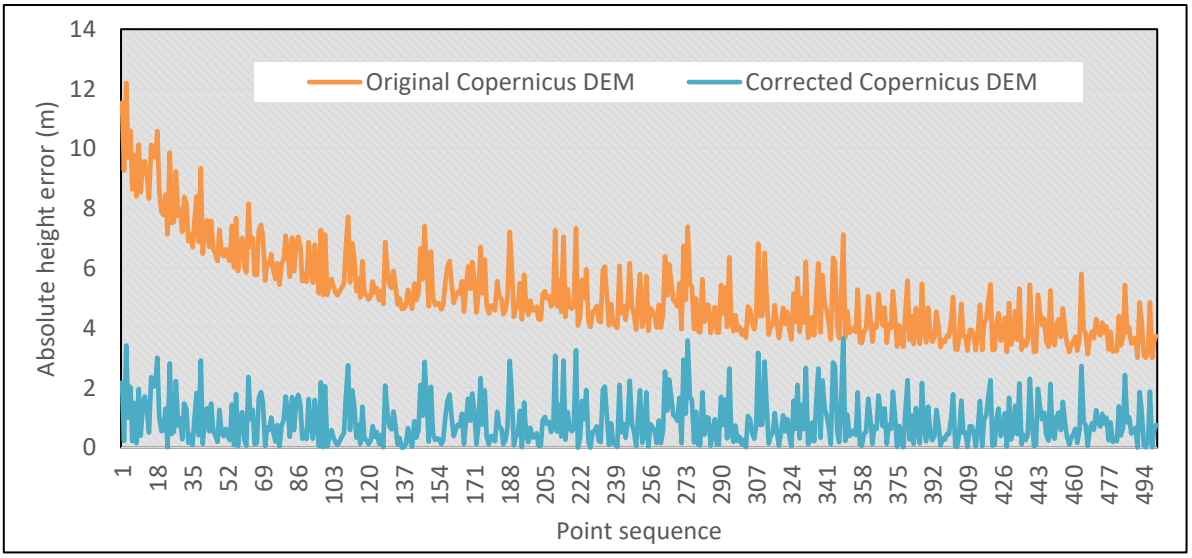


Figure 5.71: Absolute height errors at the 5th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.XGB) for 500 selected points in the urban/ industrial landscape

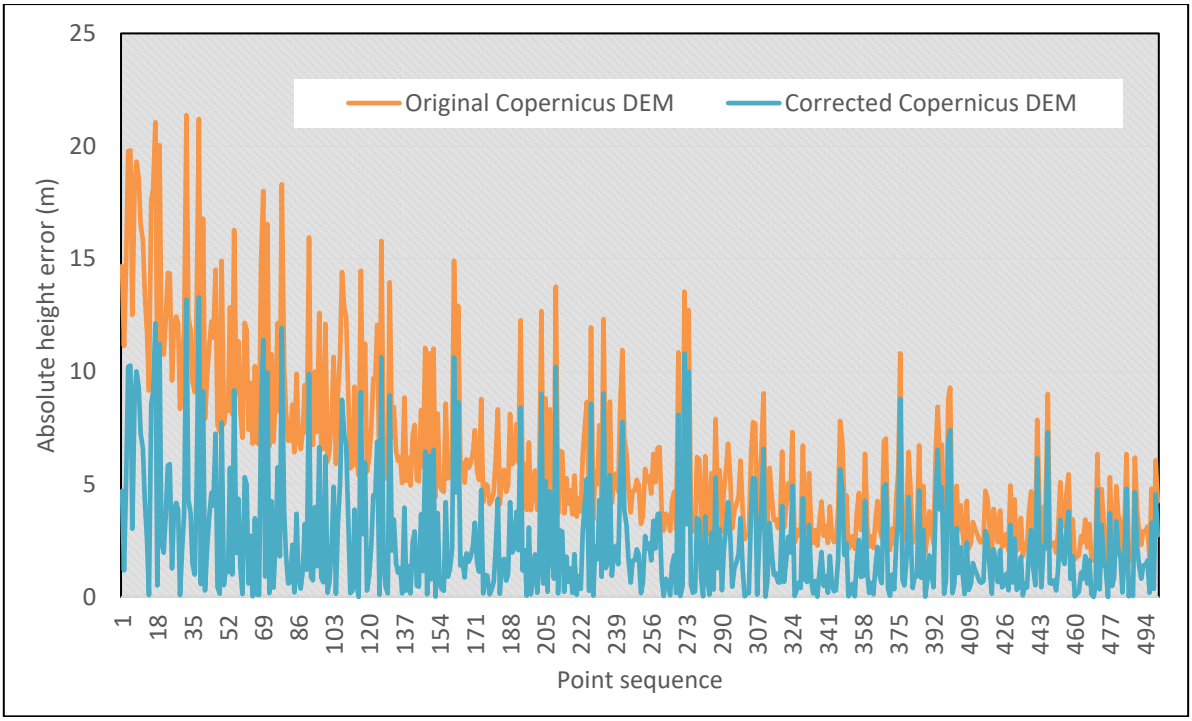


Figure 5.72: Absolute height errors at the 5th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the agricultural landscape

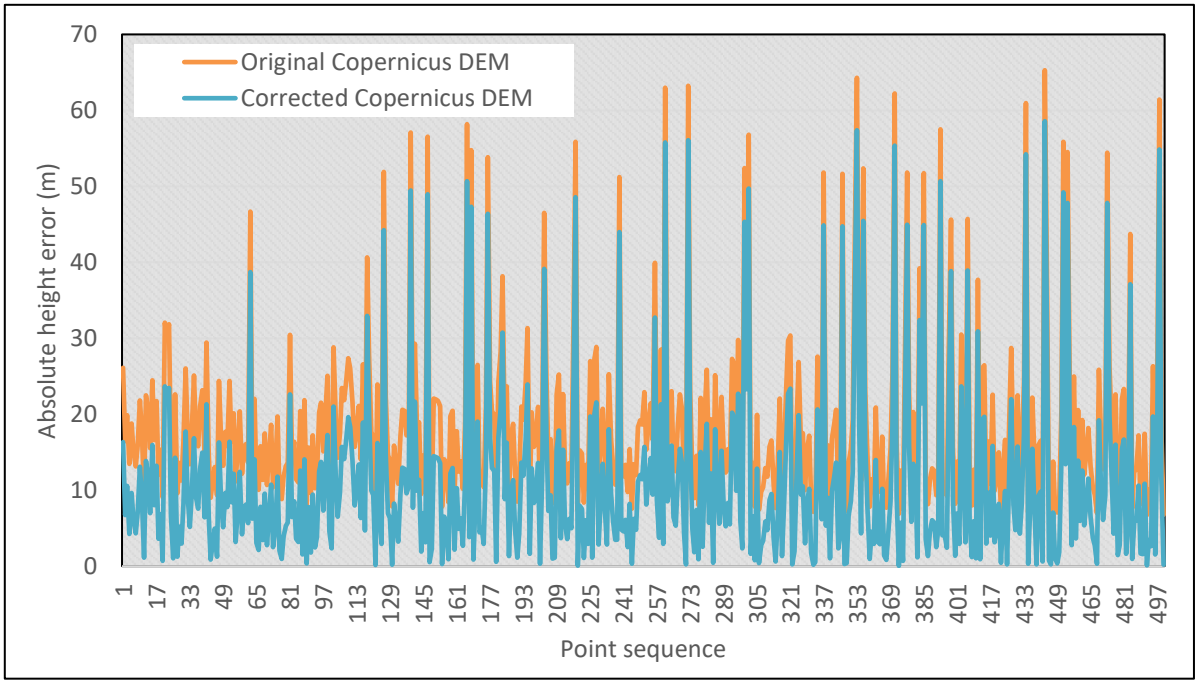


Figure 5.73: Absolute height errors at the 5th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the mountain landscape

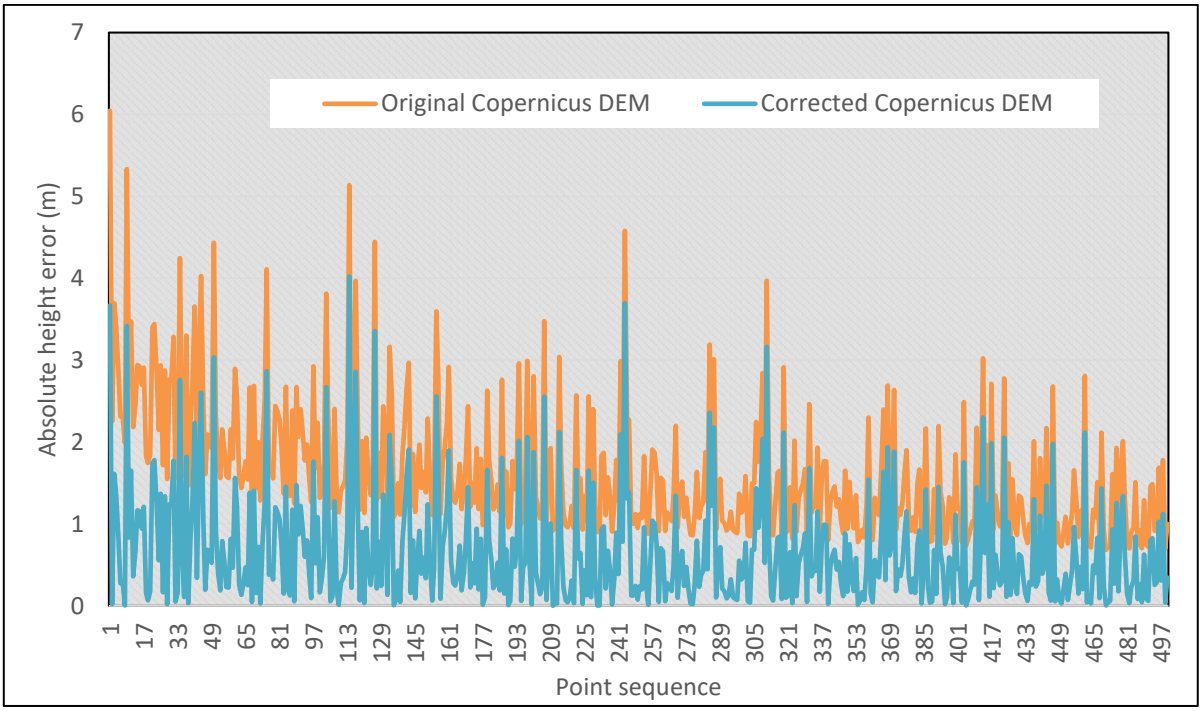


Figure 5.74: Absolute height errors at the 5th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.Cat) for 500 selected points in the peninsula landscape

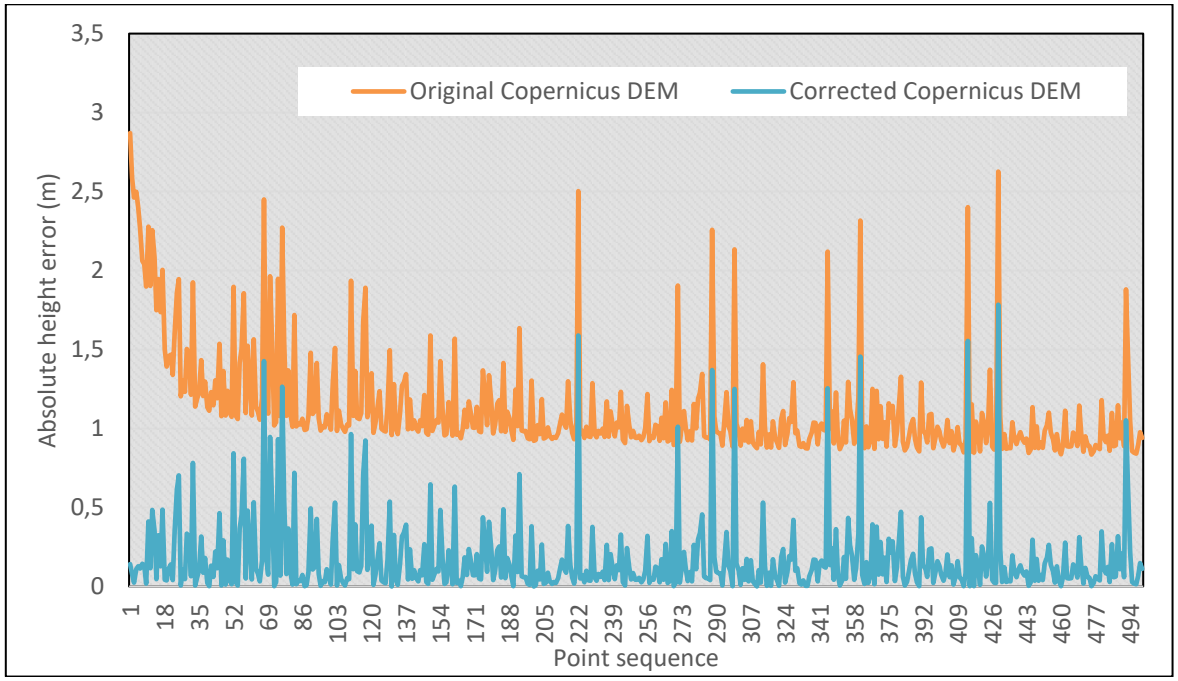


Figure 5.75: Absolute height errors at the 5th implementation site showing a comparison of the original Copernicus DEM versus the corrected Copernicus DEM (B.XGB) for 500 selected points in the grassland/shrubland landscape

Figure 5.76 and Figure 5.77 showcase comparisons of the default versus Bayesian-optimised corrected Copernicus DEMs for urban/industrial and grassland/shrubland respectively.

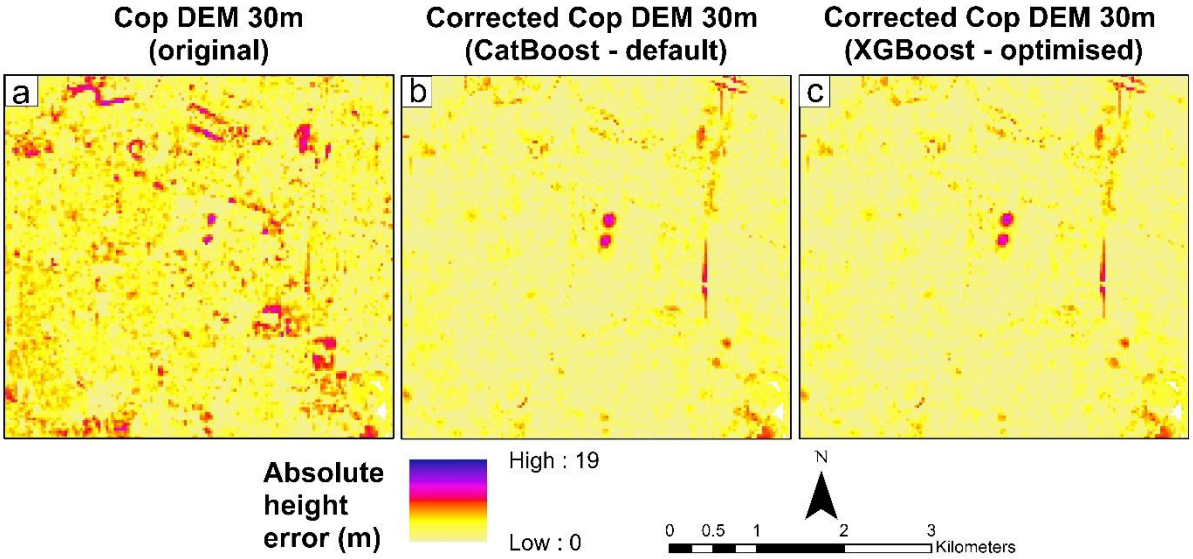


Figure 5.76: Comparison of default versus Bayesian-optimised Copernicus DEM correction at the 5th implementation site – urban/industrial

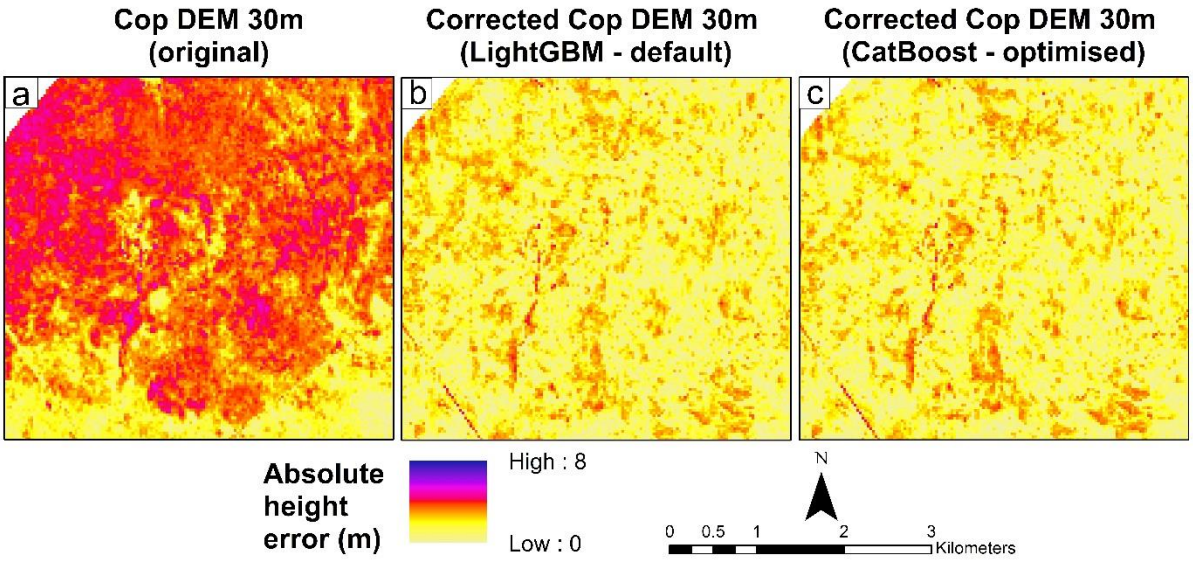


Figure 5.77: Comparison of default versus Bayesian-optimised Copernicus DEM correction at the 4th implementation site – grassland/shrubland

Generally, the default and Bayesian-optimised versions of all three models resulted in significant error reductions in the original Copernicus DEM. However, the optimised model outperformed the default models in most instances and landscapes. For example, with default hyperparameters in the urban/industrial landscape (at the 4th implementation site), the MAE reduction was in the range of 56.3% (XGBoost) to 56.9% (LightGBM). With optimisation, the improvement was increased to 57.1% with LightGBM. However at the 5th implementation site, the default CatBoost model outperformed the optimised version in the mountain

landscape, reducing the MAE from 4.848 m to 4.064 m and the RMSE from 10.237 m to 9.187 m. Similarly in the agricultural lands (at the 5th implementation site), the highest reduction of 23.9% in MAE from 0.420 m to 0.320 m was achieved by the default CatBoost model. In both implementation site, both default and optimised models achieved significant improvements.

5.4.7 Explainability of model outcomes

There is limited information on the interdependence between terrain and land cover parameters used in DEM enhancement, including their influence on the predicted targets (elevation error). To understand these interactions, SHAP (Shapley Additive Explanations) is adopted in this research as an excellent measure to explain individual predictions of the optimised models.

1. Urban/industrial

The urban/industrial SHAP plots in Figure 5.78 and Figure 5.79 provide insights on how the values of the features (predictor/input variables) influence the model predictions of DEM error. There is a tendency for higher values of the input variables to affect the predictability of the model in a positive direction. Similarly, lower feature values impact the model predictability in the negative direction. In the urban/industrial landscape, six features are found to be dominant or most influential: TPI, TST, urban footprints, tree cover, VRM and TRI. In both directions (positive and negative), the extent of SHAP values for TPI are nearly equal. This suggests that TPI could have a near-equal impact on the model predictability in both directions, and should be dealt with carefully. Thus, TPI is more influential to the predictability of the model, than the other features as shown in Figure 5.79.

Landscapes which exhibit high degree of terrain ruggedness are characterised by very high or very low TPI values, and potentially higher positive or negative SHAP values. This is a possible explanation for the dominant influence of TPI in the urban/industrial landscape. However, most of the other input variables have more positive impact than negative on the predictability of the model (Figure 5.79). Slope and aspect which are very popular first order terrain parameters are the least influential in the model predictability within the urban/industrial area.

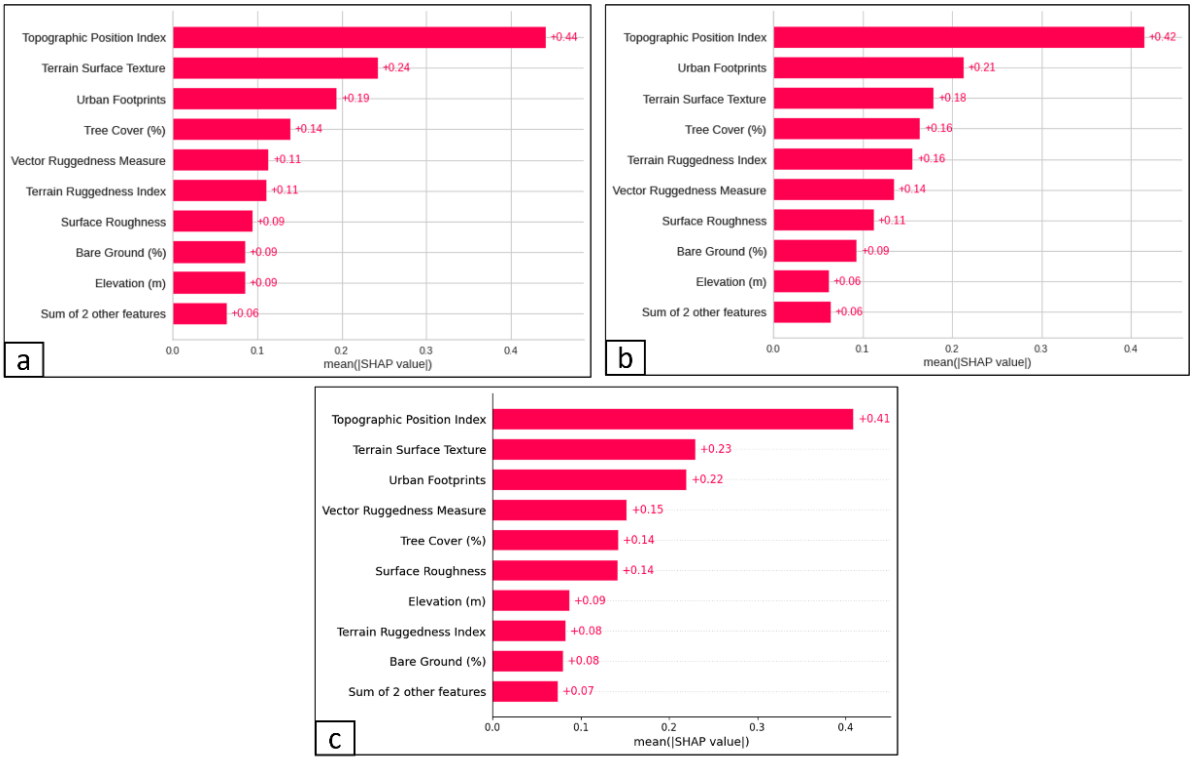


Figure 5.78: Mean SHAP values and feature importance for Bayesian-optimised prediction of Copernicus elevation error in the urban/industrial landscape, (a) XGBoost (b) LightGBM (c) CatBoost

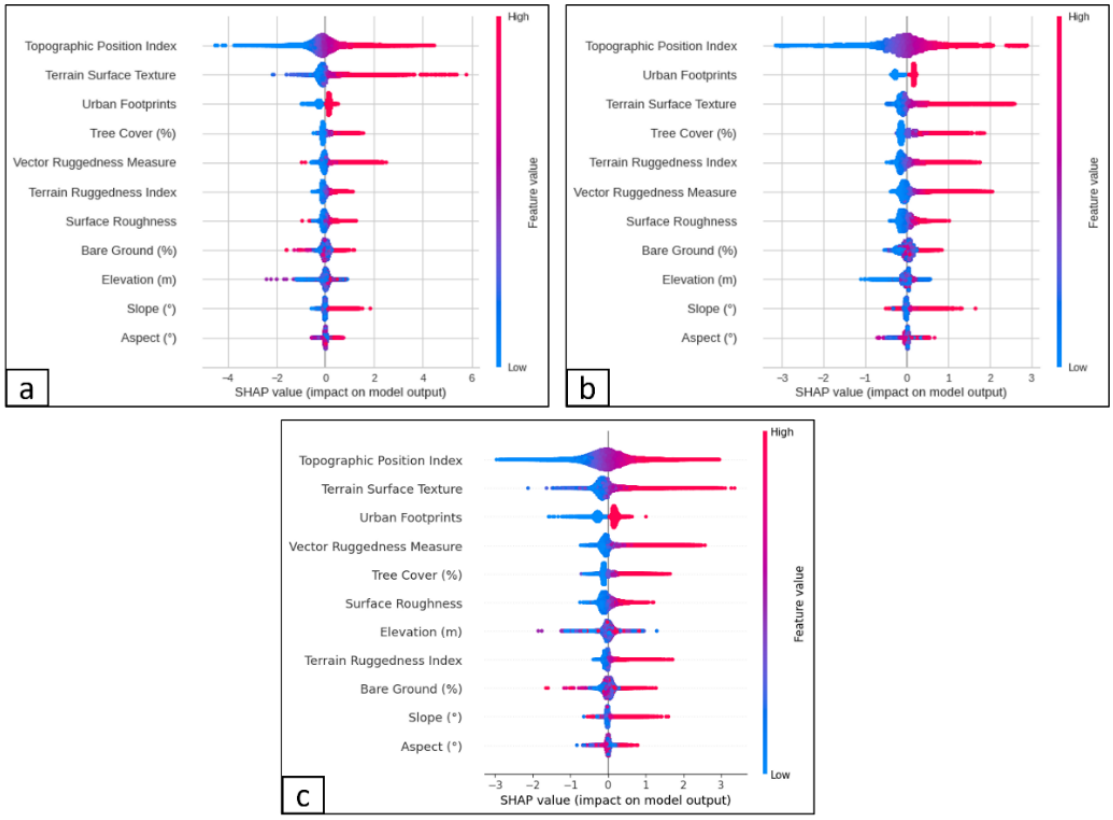


Figure 5.79: SHAP values and directional impact of features in Bayesian-optimised prediction of Copernicus elevation error in the urban/industrial area, (a) XGBoost (b) LightGBM (c) CatBoost

In many cities, buildings are constructed on land with gentle slope or gradient (Zhou et al., 2021), and possibly the slope variability could be moderate in urban areas (e.g., Figure 5.80). The residential suburbs of Cape Town extend up the slopes of the Table Mountain. Thus, the mountain slopes will start having buildings, even in some of the steeper sections, such as is evident in areas such as Camps Bay and Clifton. There are very high value properties vying for the sea views and with limited development space due to the mountains running into the oceans. Some of this may be quite unique to Cape Town, given that it has a mountainous peninsula with oceans on both sides. This restricts urban sprawl and results in housing moving up the slopes more than may occur where there is more space (e.g., the northern suburbs of Cape Town).

The likely explanation for the lower influence of slope and aspect could also be related to the characteristic cover or draping of buildings over the land surface in the urban area which makes the slope or aspect transition or changes less conspicuous when viewed from above. However, other terrain parameters could prove more influential in explaining the elevation error. For example, the VRM caters for slope and aspect variability, and incorporates 3D dispersion of vectors (Sappington et al., 2007). Moreover, VRM can differentiate among terrain types (e.g., differentiating smooth, steep hillsides from irregular terrain) and enables the treatment of terrain components as separate variables when quantifying landscapes (Sappington et al., 2007; Welty & Jeffries, 2018).

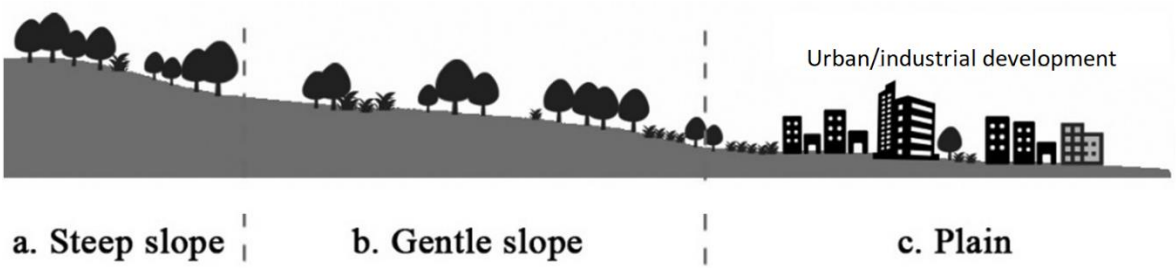


Figure 5.80: Moderate slope in urban areas, and the increasing steepness as sprawl increases upwards on mountain slopes (Adapted from Zhou et al., 2021, p. 2)

Landscapes with relatively very high/very low TPI (e.g., Figure 5.81) could be associated with increased ruggedness, and higher positive or negative SHAP values. This corroboration could explain the exceeding influence of TPI which is shown to be the highest influencer in the urban area.

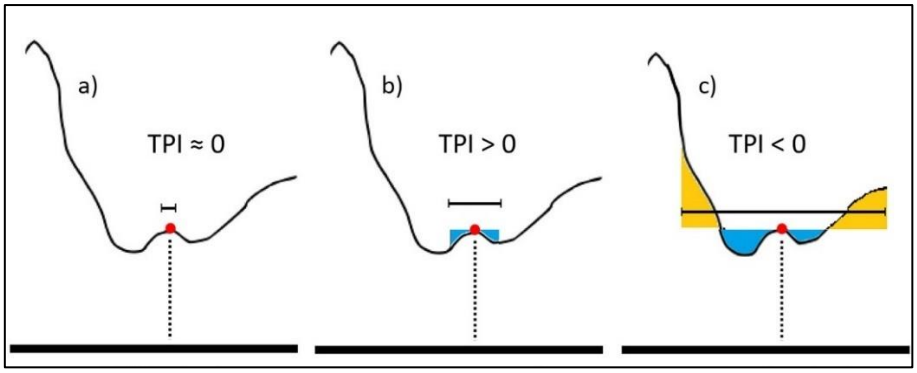


Figure 5.81: Representation of TPI for the same point (red point) at three different scales (horizontal line above red point) (Source: Adapted from Salinas-Melgoza et al., 2018, p. 7)

2. Mountain

There are visible differences in the influences of the predictor variables and observed interdependencies in the mountain landscape. For example, tree cover is more influential than TPI on the model predictability by XGBoost and LightGBM (Figure 5.82 and Figure 5.83). As the feature values of tree cover increased, its impact on the predictability of the model increased in the positive direction. The influence of TPI was similar to the scenario in the urban/industrial landscape, i.e., near-equal impact on the model predictions in both positive and negative directions. There are observed differences between the three models. For example, bare ground has the highest SHAP value in CatBoost (Figure 5.82c). However, tree cover and TPI still have a higher degree of contribution to the predictability by CatBoost, similar to their influence in XGBoost and LightGBM. The six most influential predictors in the mountainous landscape are tree cover, TPI, bare ground, elevation, slope and aspect.

Mountains are characterised by their highly steep sides and edges, and this is true of Table Mountain. This could account for the higher influence of slope and aspect (slope direction) in the mountain landscape. Moreover, there is a rich diversity of shrubland and trees on the Table Mountain, which are interspersed with rocky surfaces. Thus, the spatial variation of vegetation on the mountain is spatially correlated with the DEM error.

3. Grassland/shrubland

In the grasslands/shrublands (Figure 5.84 and Figure 5.85), the six most influential features in XGBoost are tree cover, elevation, bare ground cover, TPI, TST and VRM. However in LightGBM, aspect displaces VRM. CatBoost prioritised bare ground cover, tree cover, TPI, elevation, aspect and TST which are shown to be its six most influential features.

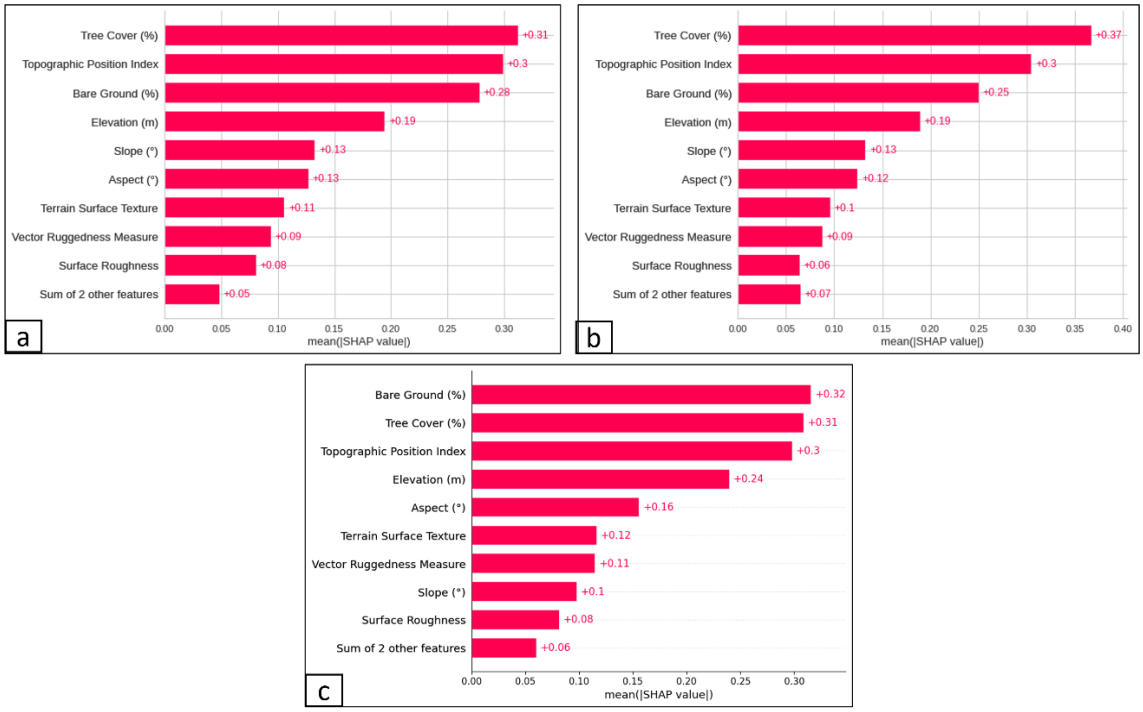


Figure 5.82: Mean SHAP values and feature importance for Bayesian-optimised prediction of Copernicus elevation error in the mountain landscape, (a) XGBoost (b) LightGBM (c) CatBoost

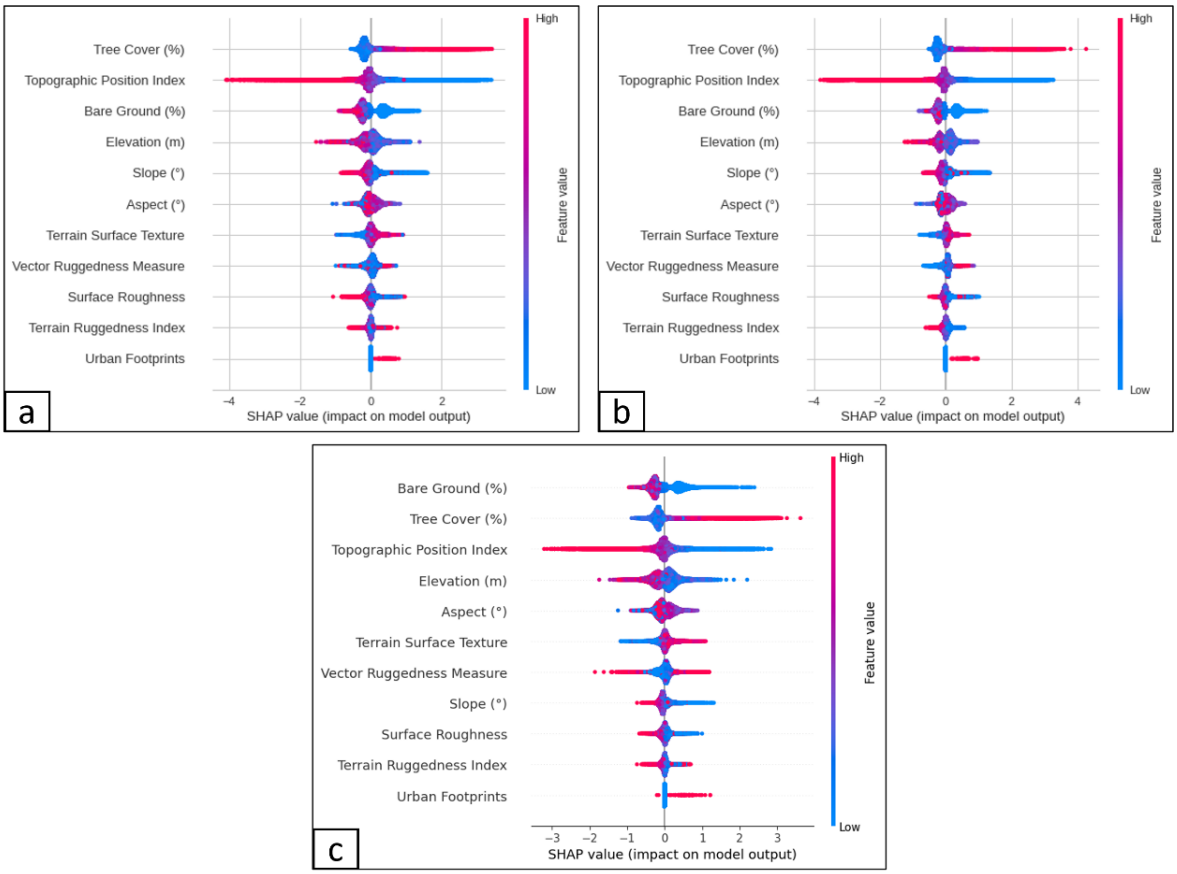


Figure 5.83: SHAP values and directional impact of features in Bayesian-optimised prediction of Copernicus elevation error in the mountain landscape, (a) XGBoost (b) LightGBM (c) CatBoost

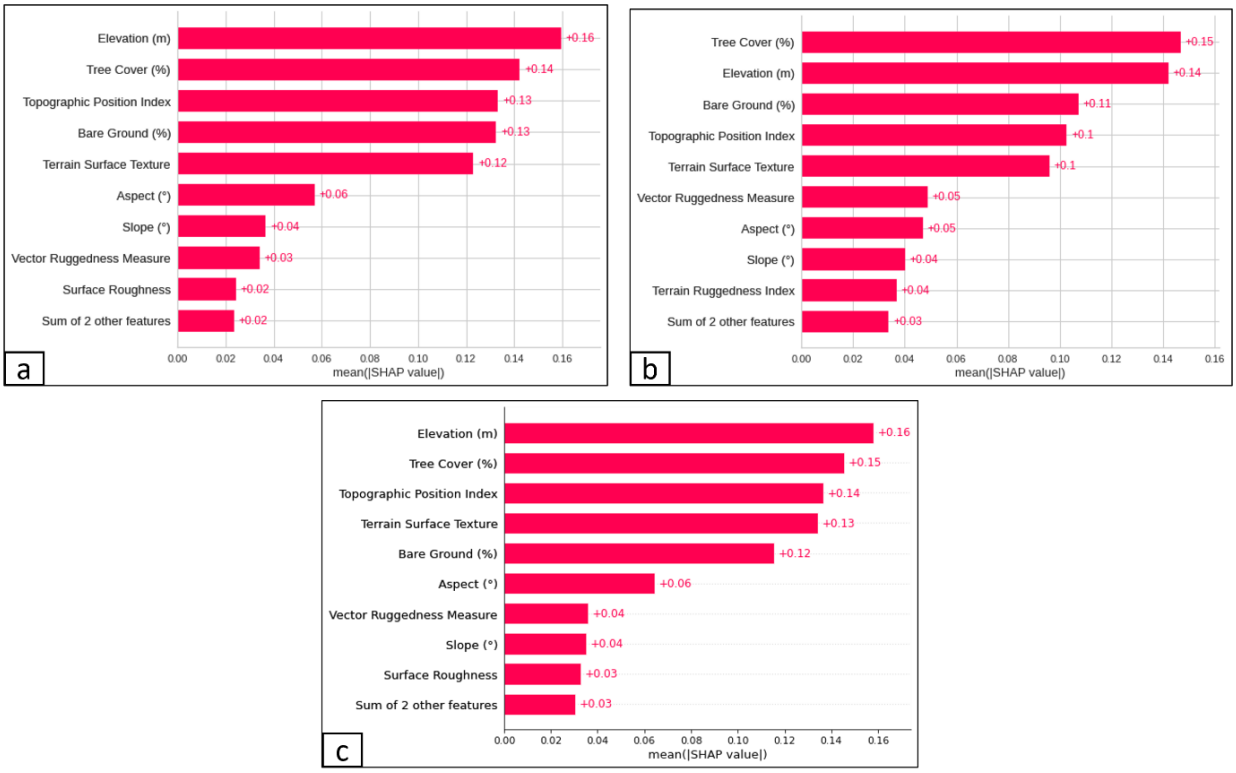


Figure 5.84: Mean SHAP values and feature importance for Bayesian-optimised prediction of Copernicus elevation error in the grassland/shrubland landscape, (a) XGBoost (b) LightGBM (c) CatBoost

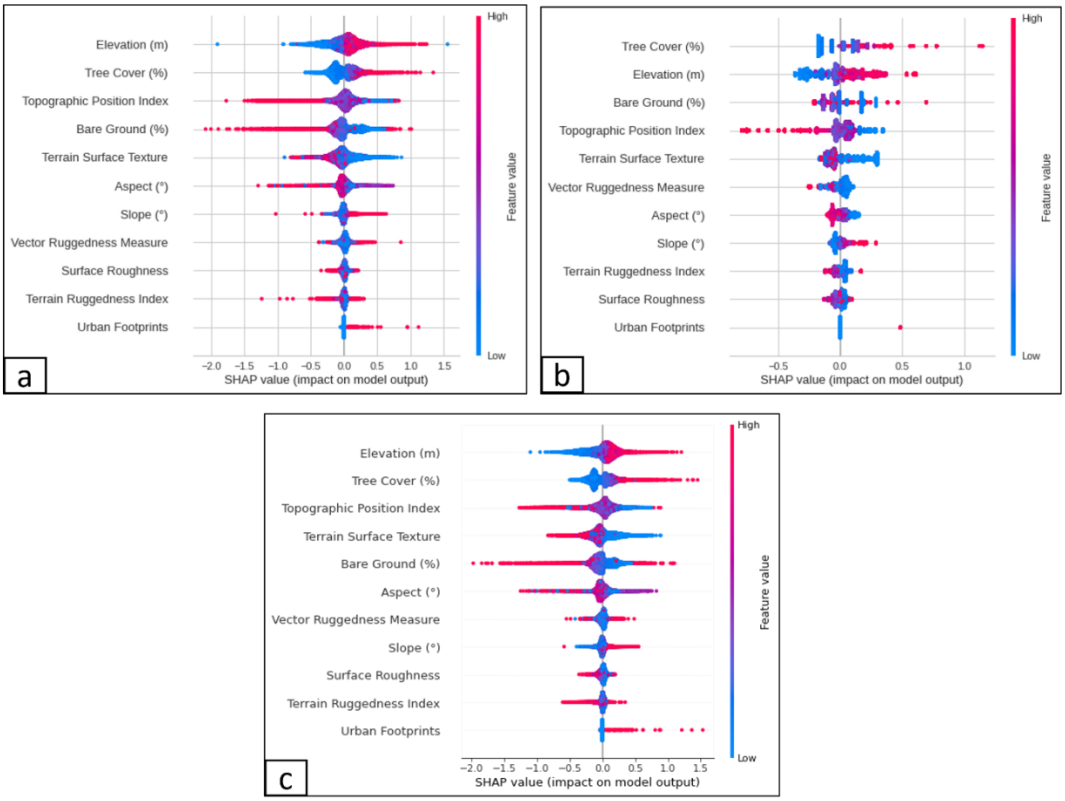


Figure 5.85: SHAP values and directional impact of features in Bayesian-optimised prediction of Copernicus elevation error in the grassland/shrubland landscape, (a) XGBoost (b) LightGBM (c) CatBoost

In the interpretation of SHAP, Ortega (2020) advises that different tree-based algorithms can have similar results, Shapley values do not necessarily imply causation, and that the correlation between variables should be factored in making conclusions.

5.4.8 Discussion of GBDT-based DEM enhancement results

The sequential DEM enhancement scheme is based on gradient boosted decision trees (GBDTs) which have emerged as a winning solution in the machine learning community. Despite this reality, the remote sensing community is yet to harness their full potential. The presented analysis is a practical demonstration of their processing and computational efficiencies, and accuracy expectations of the three most popular GBDT algorithms in the machine learning community, XGBoost, LightGBM and CatBoost. The essence of the enhancement scheme is not on the specific capabilities of the tested algorithms, but to present a variety of implementation strategies depending on user requirements/terms of reference such as computational resources, prediction accuracy and hyperparameter setup.

For several years after its release, XGBoost dominated in machine learning challenges and was regarded as a game-changer for classification and regression tasks delivering optimal speed and performance. In the results, the more recent LightGBM proved to be several times faster than XGBoost and delivered competitive results. CatBoost emerged as the best option for yielding the highest accuracy at most sites. However, for faster processing speed and when computing resources are limited, users may opt for LightGBM which provides extremely fast training speed, faster than XGBoost and CatBoost, and presents competitive results. In most instances, the default version of LightGBM surpassed the default CatBoost in the reduction of AW3D DEM RMSE at the 5th implementation site. CatBoost emerged as the overall best for delivering high prediction accuracy, followed by LightGBM and XGBoost.

This work also presents a practical analysis of how the default and hyperparameter-optimised variants of the GBDTs work in terms of training speed, prediction accuracy and hyper-parameter setup. To the best of the authors knowledge, none of the previous studies on ML-based DEM enhancement have provided such a comprehensive assessment of GBDTs despite their appealing properties. Hyperparameter tuning can be very time and resource intensive. In this study, it is estimated that the optimisation process consumed over 99.9% of the computational effort in the training of all three algorithms. However, the optimisation and training times can be significantly reduced by trimming down the search space.

Despite the observed gains in the tuned models, the time and effort required for extensive hyperparameter tuning must be taken into consideration. Depending on the user requirements, terms of reference and applications, the default versions could provide competitive performance without the computational burden of tuning. Finally, the DEM enhancement scheme presents the user with some flexibility in the choice of the GBDT algorithm to adopt, depending on the requirements.

5.5 Results of Performance Evaluation and Comparative Assessments

5.5.1 Terrain analysis and 3D visualisation

Contour lines are used for encoding elevation information on topographic maps (Atit et al., 2016). In this section, cartographic contours interpolated from the DEMs are compared. The first evaluation involved comparison of 10 m interpolated (and smoothened) contours, 3D terrain surfaces and shaded relief maps from the LiDAR DEM, original Copernicus DEM, corrected Copernicus DEM and South Africa National DEM. Figure 5.86 presents contour maps interpolated from the original Copernicus DEM, the corrected Copernicus DEM and LiDAR, whereas Figure 5.87 presents contour maps interpolated from the South Africa National DEM, the corrected Copernicus DEM and the LiDAR DEM.

The results indicate that contours generated from the enhanced DEMs could outperform the original DEM in matching the geometry of base contours generated from airborne LiDAR, likely due to the reduction in above-ground terrain offsets achieved during the correction. The enhanced DEMs could also yield contours with better cartographic quality by jettisoning redundant vertices which are a burden for efficient contour generalisation in topographic maps. The analysis also shows the capability of the DEM enhancement scheme to yield contours that closely match the representation by the authoritative SA National DEM and LiDAR DEM. However, the causes of some planimetry differences between contours generated from the SA National DEM and other DEMs are beyond the scope of this comparison. The 3D surfaces and shaded relief maps (Figure 5.88 and Figure 5.89) show evidence of terrain offsets, irregularities and bumps in the original Copernicus DEM along the Diep River catchment which was nearly flattened out or diminished in the corrected Copernicus DEM. The assessment indicates that the enhancement can improve the detection of physical landforms on topographic maps. This points to the suitability of the enhanced satellite-derived DEMs as a potential substitute or alternative for very costly airborne topographic mapping.

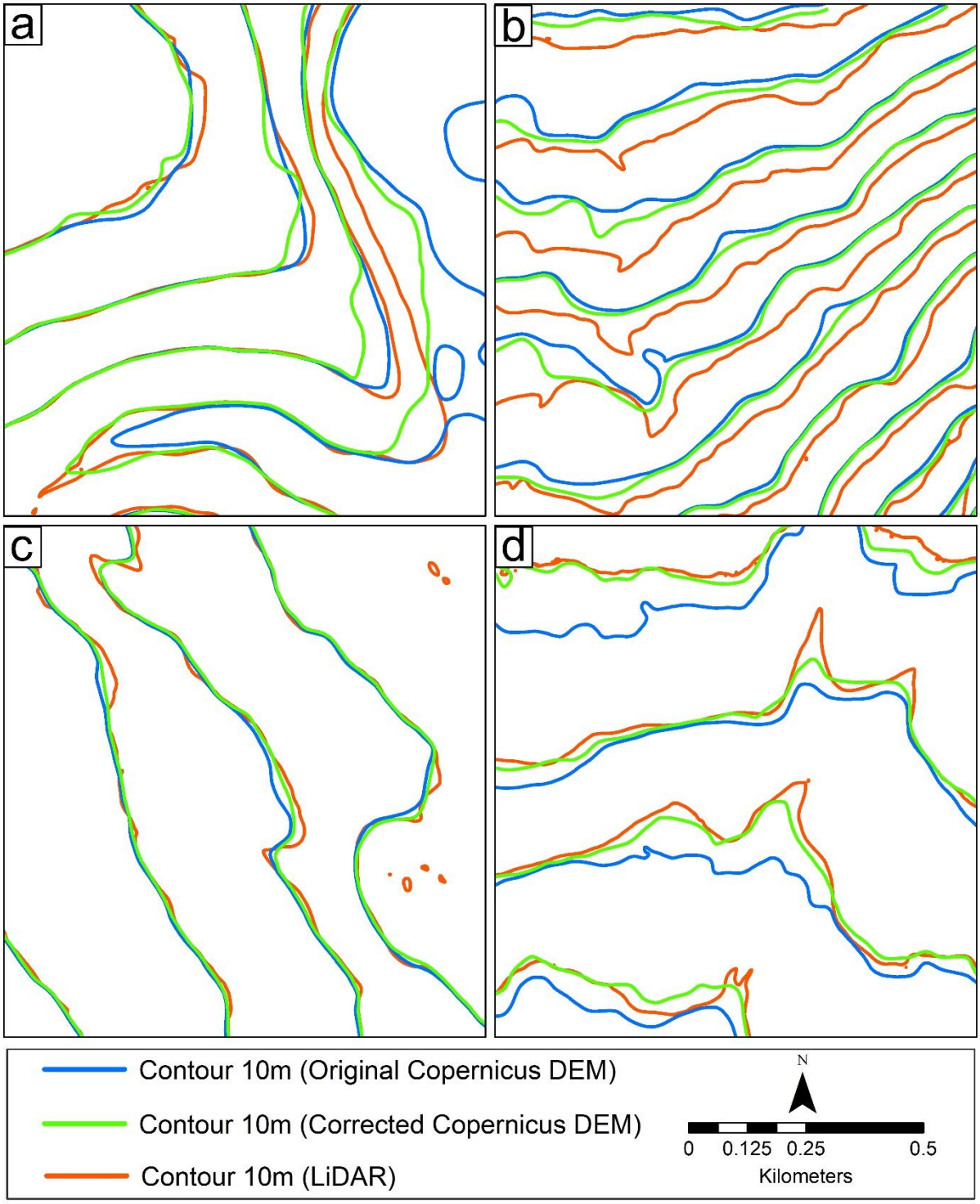


Figure 5.86: Contour maps interpolated from the 30 m original Copernicus DEM, the 30 m corrected Copernicus DEM and 2 m LiDAR, (a) Agricultural land, 4th implementation site (b) Mountain, 4th implementation site (c) Peninsula, 5th implementation site (d) Grassland/ shrubland, 4th implementation site

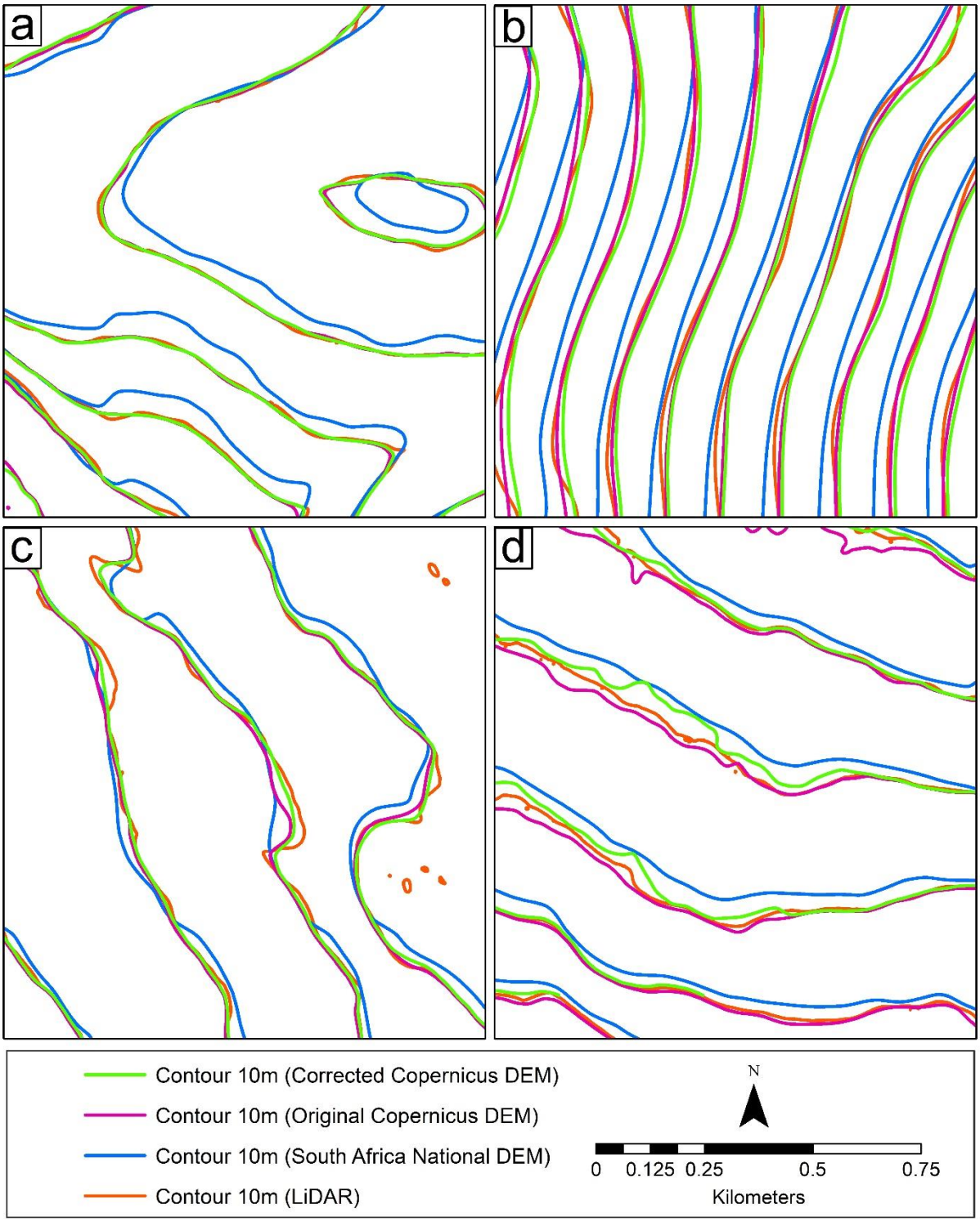


Figure 5.87: Contour maps interpolated from the 30 m corrected and original Copernicus DEMs, the 25 m South Africa National DEM and the 2 m LiDAR DEM, (a) Agricultural land, 4th implementation site (b) Mountain, 4th implementation site (c) Peninsula, 5th implementation site (d) Grassland/ shrubland, 4th implementation site

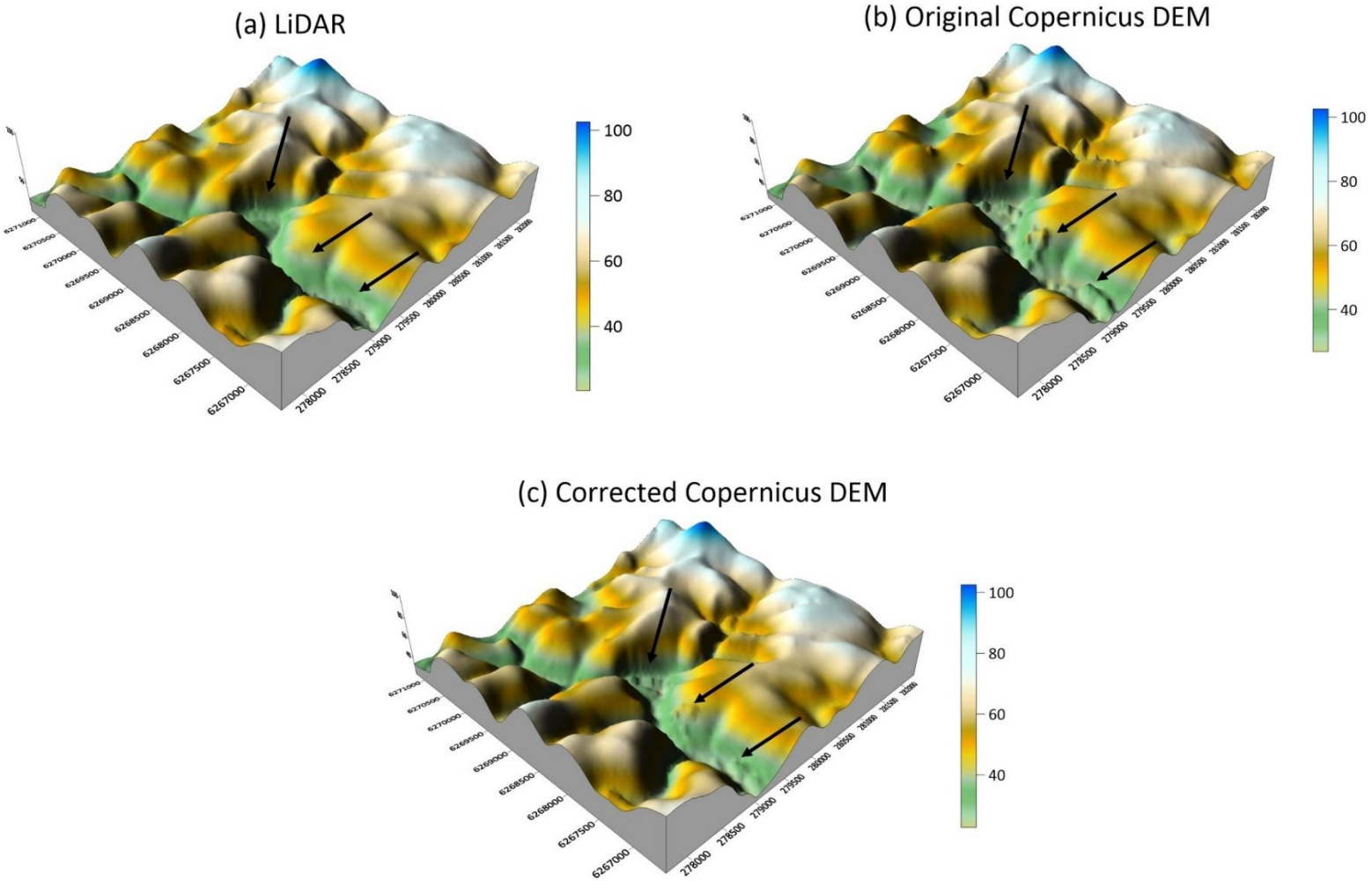


Figure 5.88: Comparison of 3D terrain representation of the agricultural land generated from (a) 2 m LiDAR (b) 30 m Original Copernicus DEM and (c) 30 m Corrected Copernicus DEM, at the 4th implementation site

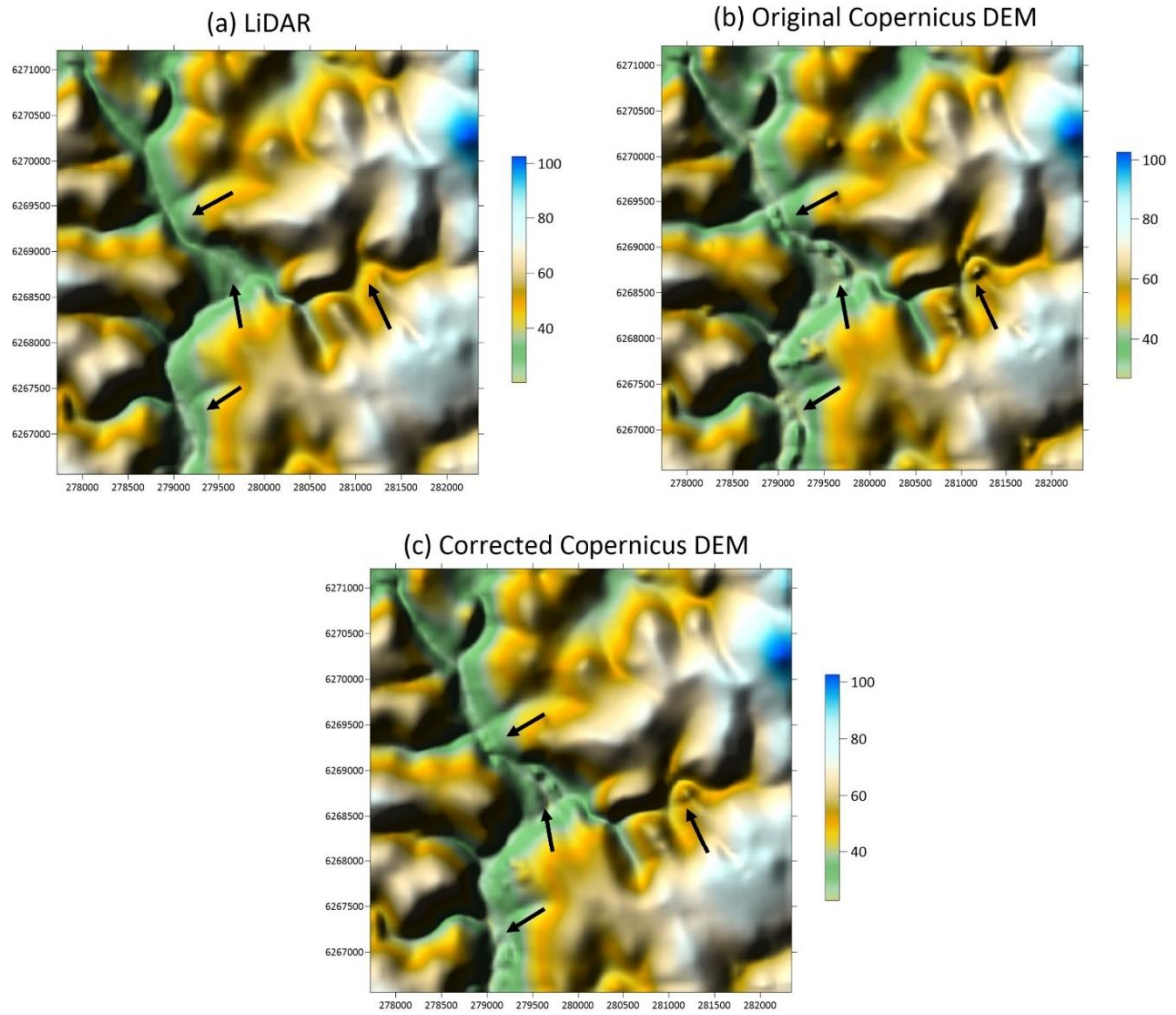


Figure 5.89: Comparison of shaded relief maps of the agricultural land generated from (a) 2 m LiDAR (b) 30 m Original Copernicus DEM and (c) 30 m Corrected Copernicus DEM, at the 4th implementation site

5.5.2 Delineation and mapping of lineaments

This comparison evaluates the potential of the global DEMs to match the density of lineaments derived from the LiDAR DEM. For the lineament mapping, the 5 m version of the LiDAR DEM was used, and the Copernicus DEMs were resampled to 5 m for consistency. The lineaments were extracted over the Table Mountain landscape through analytical shading of the DEMs (Figure 5.90) followed by edge detection, thresholding and curve detection.

Comparison of the delineated lineaments is achieved using lineament maps, lineament density maps and rose diagrams showing the lineament orientations. The analysis shows that the LiDAR DEM was the most sensitive for delineating the surficial and rectilinear terrain patterns which manifested as lineaments on the Table Mountain. The corrected Copernicus DEM surpassed the sensitivity of the original Copernicus DEM by outputting a higher magnitude (Figure 5.91) and a denser configuration of lineaments (Figure 5.92). The rose diagrams

(Figure 5.93) suggest that the corrected DEMs can achieve better homogeneity in the lineament orientation that better matches the base lineaments from LiDAR, a likely effect of the diminution of terrain offsets during the height correction.

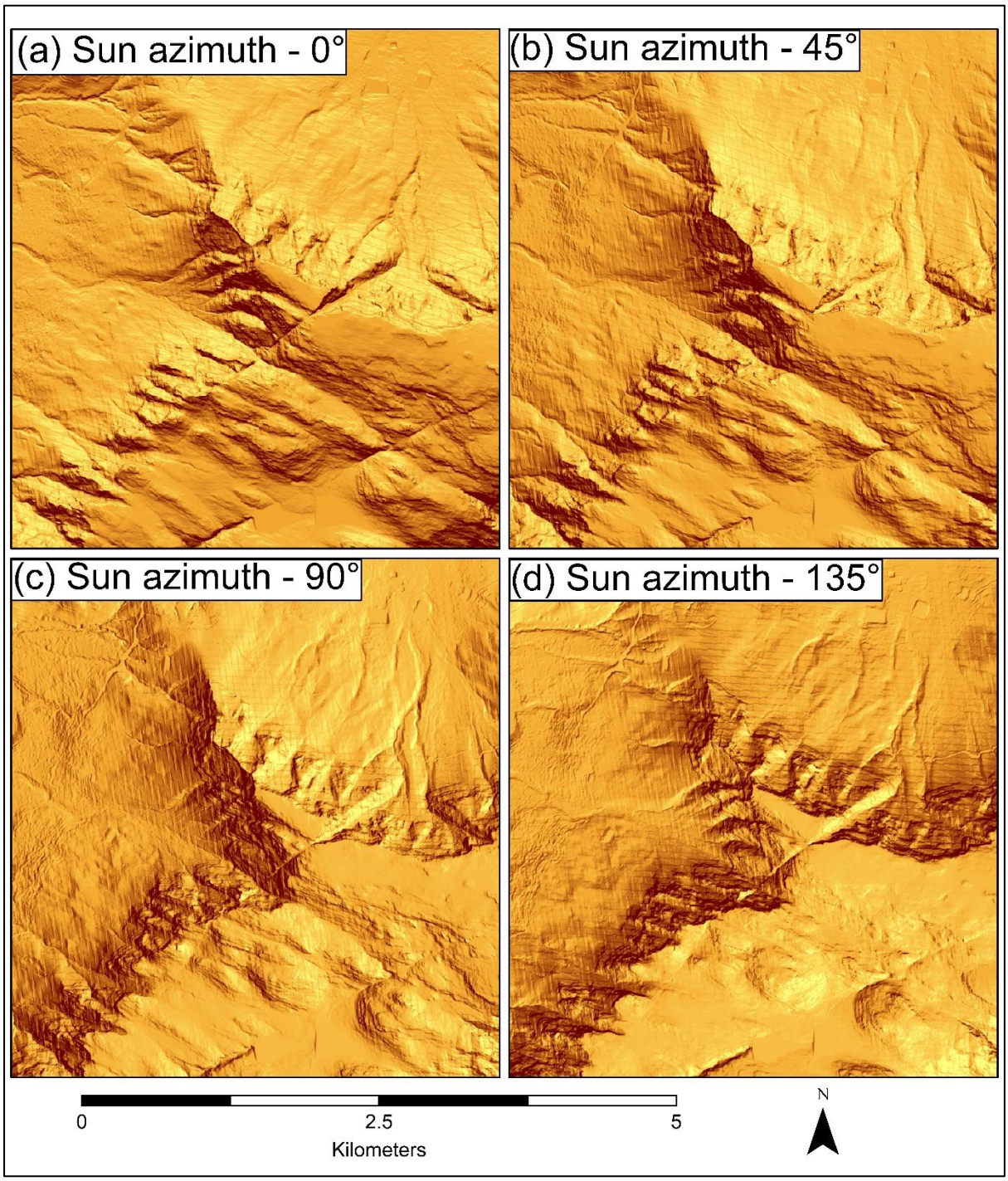


Figure 5.90: Analytical shading of the mountain landscape at the 4th implementation site, generated from 5 m airborne LiDAR-derived DEM at different sun azimuths

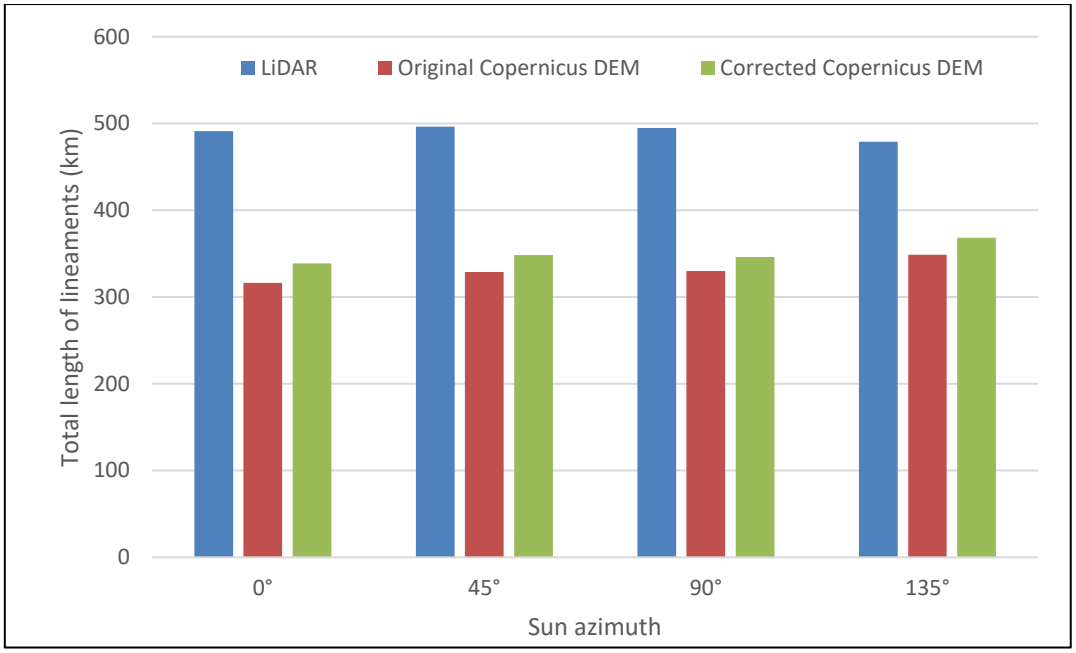


Figure 5.91: Visual comparison of the total length of lineaments generated from shaded relief maps of LiDAR DEM, the original Copernicus DEM and the corrected Copernicus DEM, on mountain landscape at the 4th implementation site

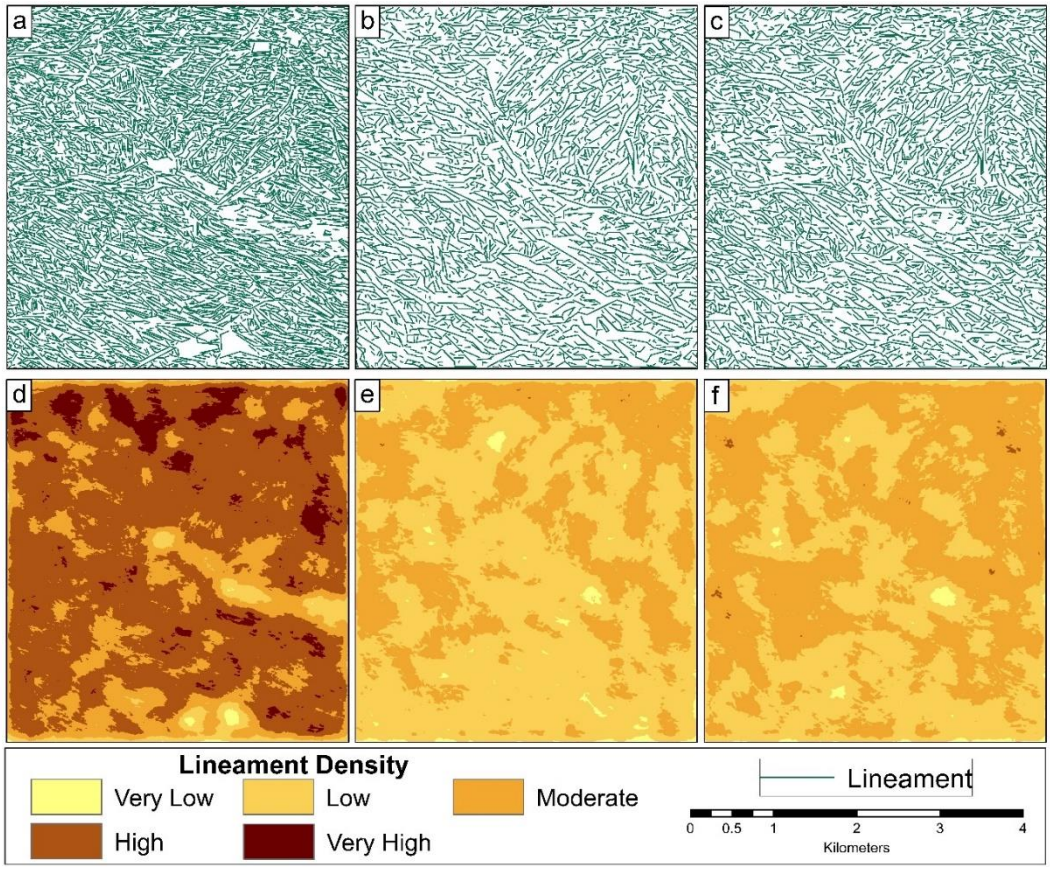


Figure 5.92: Comparison of lineaments and lineament densities in the mountain landscape (4th implementation site) generated at 0° sun azimuth – (a) lineaments from LiDAR DEM (b) lineaments from Original Copernicus DEM (c) lineaments from corrected Copernicus DEM (d) lineament density from LiDAR DEM (e) lineament density from Original Copernicus DEM (f) lineament density from Corrected Copernicus DEM

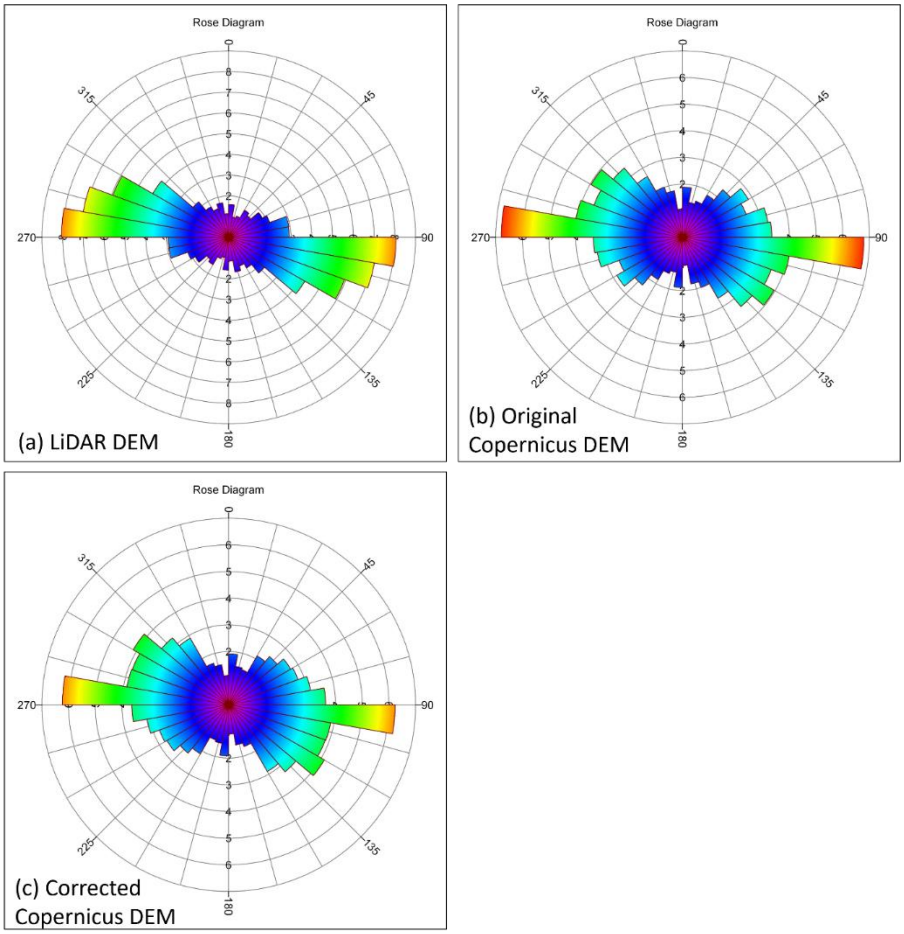


Figure 5.93: Comparison of rose diagrams showing the lineament orientation in the mountain landscape (4th implementation site) generated at 0° sun azimuth – (a) LiDAR DEM (b) Original Copernicus DEM (c) Corrected Copernicus DEM

5.5.3 Comparison with multiple linear regression

The correlation analysis and VIF tests detected three variables with high collinearity – slope, TRI and surface roughness. Figure 5.94 and Figure 5.95 show correlation heatmaps of the terrain parameters from AW3D and Copernicus DEMs respectively. This relationship was corroborated by the VIF tests in which slope, TRI and surface roughness emerged with VIFs greater than 10 (Table 5.33). Consequently, TRI and surface roughness (SR) were excluded from the regression analysis.

The correlation analysis of terrain parameters derived from AW3D DEM (Figure 5.94) shows that generally, TST and VRM, TRI and VRM, surface roughness and VRM, slope and VRM, surface roughness and elevation, TRI and elevation, and TST and elevation have a moderately strong correlation. There are strong correlations between TRI and TST, surface roughness and TST, slope and TST, and slope and elevation. Slope, surface roughness and TRI are very strongly correlated.

Figure 5.94: Correlation heatmap of AW3D DEM terrain parameters

In Copernicus DEM (Figure 5.95), TST and VRM, TRI and VRM, surface roughness and TST, slope and TST, surface roughness and VRM, slope and VRM, and TST and elevation have a moderately strong correlation. There are strong correlations between TRI and TST, surface roughness and elevation, TRI and elevation, and slope and elevation. Slope, surface roughness and TRI are very strongly correlated.

Figure 5.95: Correlation heatmap of Copernicus DEM terrain parameters

Table 5.33: The calculated VIF values for Copernicus and AW3D

Terrain Parameter	Copernicus	AW3D
Elevation	2.91	2.89
Slope	77.71	60.90
Aspect	1.78	1.87
Surface roughness	120.14	100.65
Topographic position index	1.02	1.02
Terrain roughness index	155.06	119.04
Terrain surface texture	2.78	3.75
Vector ruggedness measure	1.84	2.11

Thus, after excluding surface roughness and TRI, the elevation error is given as follows:

$$\text{Elevation error} = \beta_0 + (\beta_1 * \text{Elevation}) + (\beta_2 * \text{Slope}) + (\beta_3 * \text{Aspect}) + (\beta_4 * \text{TPI}) + (\beta_5 * \text{TST}) + (\beta_6 * \text{VRM}) + (\beta_7 * \text{Urban Footprint value}) + (\beta_8 * \% \text{ Tree cover}) + (\beta_9 * \% \text{ Bare ground}) \quad (5.1)$$

Table 5.34 and Table 5.35 present the derived coefficients of the input variables for AW3D and Copernicus DEMs respectively.

Table 5.34: Derived coefficients of the input variables for statistical correction of AW3D DEM

Parameter	Urban/ industrial	Agricultural	Mountain	Peninsula	Grassland/ shrubland
β_0 (Intercept)	1.581000	2.343000	1.628000	2.050000	2.952000
β_1 (Elevation)	0.008281	0.002053	-0.000882	-0.000433	0.002648
β_2 (Slope)	0.062240	-0.071080	0.007549	0.027990	0.024900
β_3 (Aspect)	-0.000122	-0.000083	0.000166	-0.001174	-0.000356
β_4 (TPI)	0.956300	0.722000	0.206300	0.374900	0.078050
β_5 (TST)	0.069460	-0.001601	0.021490	0.025380	0.002957
β_6 (VRM)	225.000000	109.400000	10.670000	-51.500000	-27.220000
β_7 (Urban Footprints)	0.002638	0.003683	0.004503	0.003504	0.005733
β_8 (Tree Cover)	0.012120	0.034210	0.032130	0.017710	0.006124
β_9 (Bare ground)	0.011080	0.002033	0.012040	-0.000188	-0.002766

Table 5.35: Derived coefficients of the input variables for statistical correction of Copernicus DEM

Parameter	Urban/ industrial	Agricultural	Mountain	Peninsula	Grassland/ shrubland
β_0 (Intercept)	-0.105300	0.021110	0.530900	-0.213600	0.443000
β_1 (Elevation)	0.000333	-0.001242	-0.000395	-0.000599	0.003351
β_2 (Slope)	0.113100	-0.021620	-0.011540	0.019820	0.032210
β_3 (Aspect)	-0.000057	0.000206	-0.000017	0.000041	-0.000607
β_4 (TPI)	0.850800	0.334500	-0.265800	-0.290900	-0.198300
β_5 (TST)	0.056560	0.015810	0.013490	0.012780	-0.013150
β_6 (VRM)	337.600000	192.500000	4.092000	-5.504000	-27.920000
β_7 (Urban Footprints)	0.002150	0.002147	0.002822	0.002772	0.002672
β_8 (Tree Cover)	0.022500	0.041920	0.031060	0.021600	0.020710
β_9 (Bare ground)	-0.000194	0.004190	-0.006962	0.002368	-0.001589

Following the calculation of elevation error, the corrections were applied to the original DEMs to derive corrected DEMs. Consequently, the statistical-based correction achieved reasonable corrections of AW3D and Copernicus DEMs, but could not surpass the GBDT approach (Table 5.36 - Table 5.39; Figure 5.96 - Figure 5.100). The comparison proves the superiority of the GBDT-based correction in virtually all the landscapes under consideration.

Table 5.36: Comparison of statistical-based and GBDT-based correction of AW3D DEM at the 4th implementation site. The best accuracies are highlighted

Landscape	**Original AW3D DEM**		**Corrected AW3D DEM**							
			Multiple Linear Regression		**XGBoost Default**		**LightGBM Default**		**CatBoost Default**	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Urban/ industrial	4.462	4.948	0.989	1.478	0.969	1.490	**0.967**	**1.468**	0.969	1.484
Agricultural	2.483	3.036	0.977	1.514	0.910	1.361	0.910	1.374	**0.905**	**1.358**
Mountain	7.999	15.465	6.895	14.639	6.833	14.609	**6.759**	14.465	6.775	**14.440**
Peninsula	2.265	2.658	0.891	1.495	0.877	1.441	**0.873**	1.431	0.879	**1.420**
Grassland/ shrubland	4.085	4.166	0.946	1.133	0.778	0.960	0.792	0.971	**0.751**	**0.929**

Table 5.37: Comparison of statistical-based and GBDT-based correction of AW3D DEM at the 5th implementation site. The best accuracies are highlighted

Landscape	**Original AW3D DEM**		**Corrected AW3D DEM**							
			Multiple Linear Regression		**XGBoost Default**		**LightGBM Default**		**CatBoost Default**	
	MAE	**RMSE**	**MAE (m)**	**RMSE (m)**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**
Urban/ industrial	3.369	3.636	0.709	1.016	0.692	1.009	0.685	**0.990**	**0.684**	1.001
Agricultural	2.258	2.735	1.098	1.514	1.066	1.456	**1.053**	**1.438**	1.058	1.440
Mountain	7.443	11.442	6.200	9.913	5.902	9.453	5.913	9.504	**5.877**	**9.427**
Peninsula	2.297	2.604	0.845	1.259	**0.824**	1.244	0.825	**1.236**	**0.824**	1.251
Grassland/ shrubland	3.565	3.660	0.759	1.002	0.778	1.001	**0.754**	**0.982**	0.761	0.985

Table 5.38: Comparison of statistical-based and GBDT-based correction of Copernicus DEM at the 4th implementation site. The best accuracies are highlighted

Landscape	**Original Copernicus DEM**		**Corrected Copernicus DEM**													
			Multiple Linear Regression		**XGBoost**				**LightGBM**				**CatBoost**			
					Default		**Optimised**		**Default**		**Optimised**		**Default**		**Optimised**	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Urban/ industrial	1.389	1.952	0.729	1.046	0.607	0.944	0.615	0.955	0.598	0.935	**0.596**	0.948	0.604	**0.933**	0.608	0.941
Agricultural	0.456	1.466	0.413	1.119	0.352	0.932	0.355	0.937	0.347	0.940	0.463	1.330	**0.346**	0.926	**0.346**	**0.895**
Mountain	6.807	17.339	6.536	17.307	6.456	17.214	6.445	**17.179**	6.441	17.184	**6.434**	17.187	6.438	17.182	6.438	17.193
Peninsula	0.424	0.905	0.407	0.822	0.361	0.725	0.358	0.716	0.357	0.708	0.363	0.732	0.356	0.699	**0.352**	**0.693**
Grassland/ shrubland	1.357	1.520	0.646	0.785	0.472	0.592	0.475	0.603	0.457	0.570	0.544	0.663	0.459	0.571	**0.434**	**0.550**

Table 5.39: Comparison of statistical-based and GBDT-based correction of Copernicus DEM at the 5th implementation site. The best accuracies are highlighted

Landscape	**Original Copernicus DEM**		**Corrected Copernicus DEM**													
			Multiple Linear Regression		**XGBoost**				**LightGBM**				**CatBoost**			
					Default		**Optimised**		**Default**		**Optimised**		**Default**		**Optimised**	
	MAE (m)	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**	**MAE (m)**	**RMSE (m)**
Urban/ industrial	1.087	1.479	0.553	0.824	0.518	0.787	**0.509**	**0.781**	0.519	0.789	0.517	0.785	0.512	0.783	**0.509**	0.788
Agricultural	0.420	1.230	0.391	0.969	0.324	0.770	0.326	0.784	0.325	0.776	0.434	1.122	**0.320**	0.764	0.325	**0.760**
Mountain	4.848	10.237	4.358	9.604	4.092	9.213	4.127	9.270	4.084	9.233	4.141	9.321	**4.064**	**9.187**	4.080	9.206
Peninsula	0.345	0.521	0.338	0.477	0.303	0.448	0.297	0.437	0.301	0.445	0.308	0.459	0.295	0.437	**0.292**	**0.430**
Grassland/ shrubland	0.497	0.563	0.250	**0.330**	0.267	0.378	**0.242**	0.353	0.258	0.364	0.313	0.400	0.276	0.381	0.266	0.373

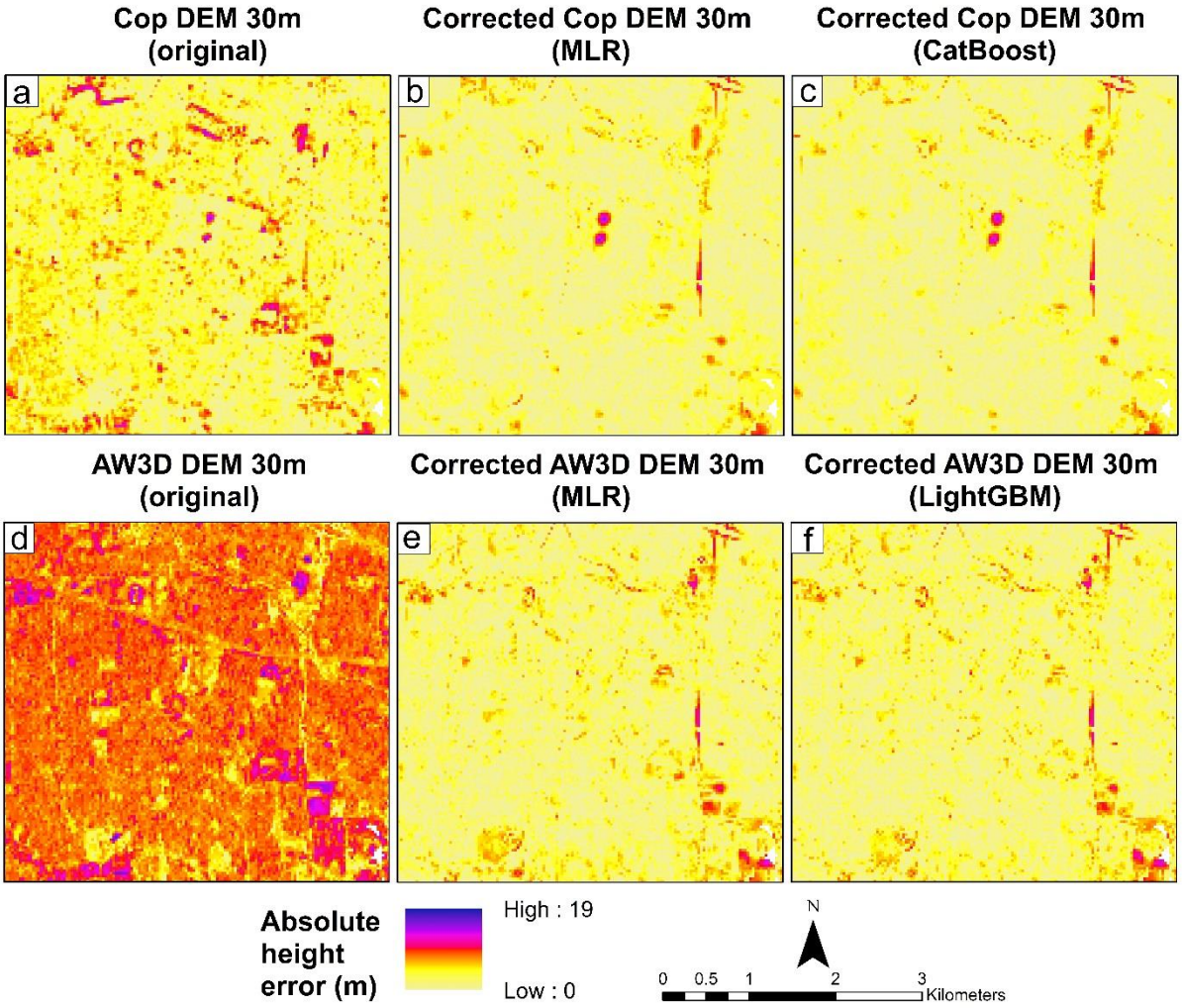


Figure 5.96: Absolute height error comparison of statistical-based and the best default GBDT-based Copernicus DEM correction at the 5th implementation site – urban/ industrial

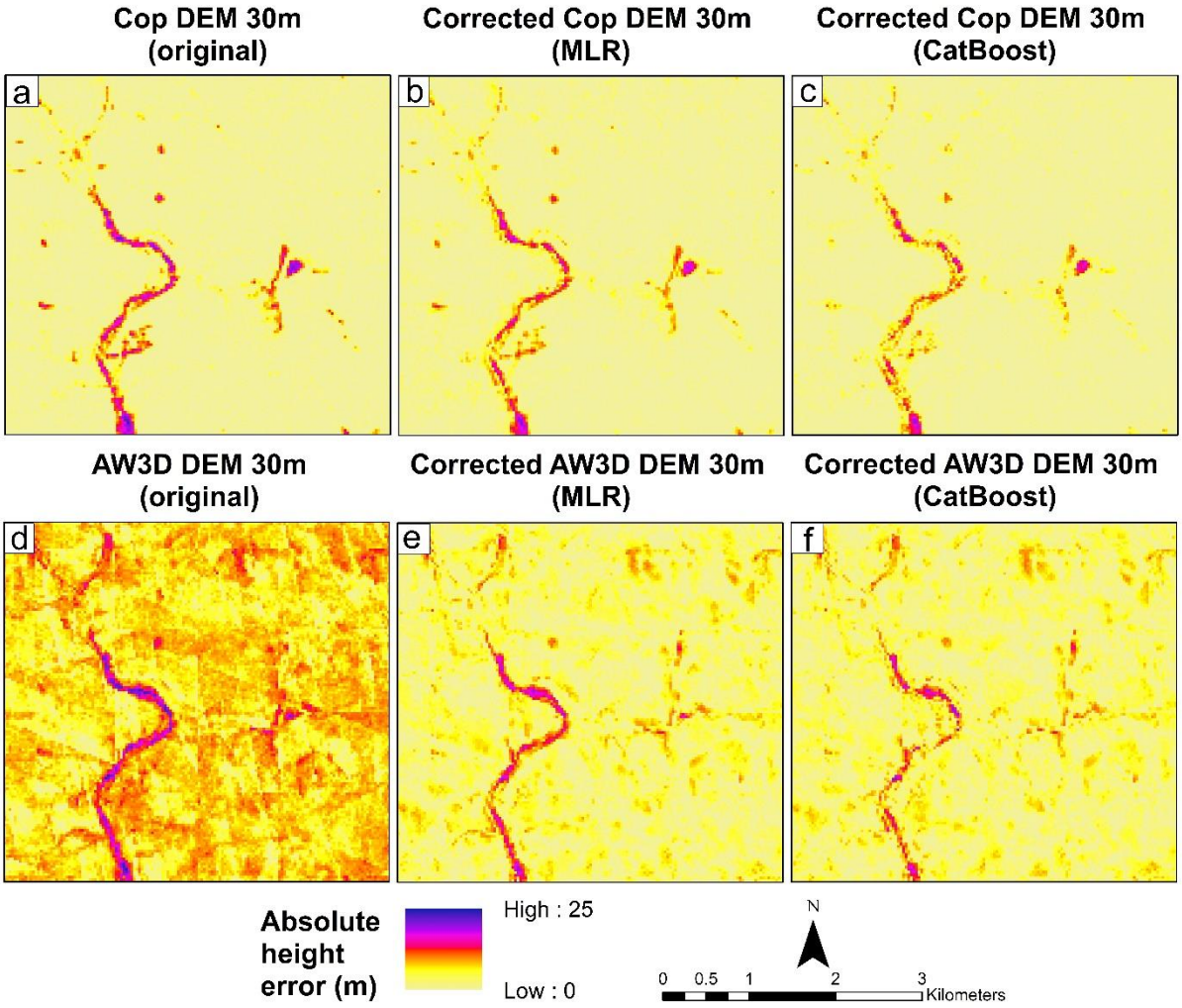


Figure 5.97: Absolute height error comparison of statistical-based and the best default GBDT-based Copernicus DEM correction at the 4th implementation site – agricultural land

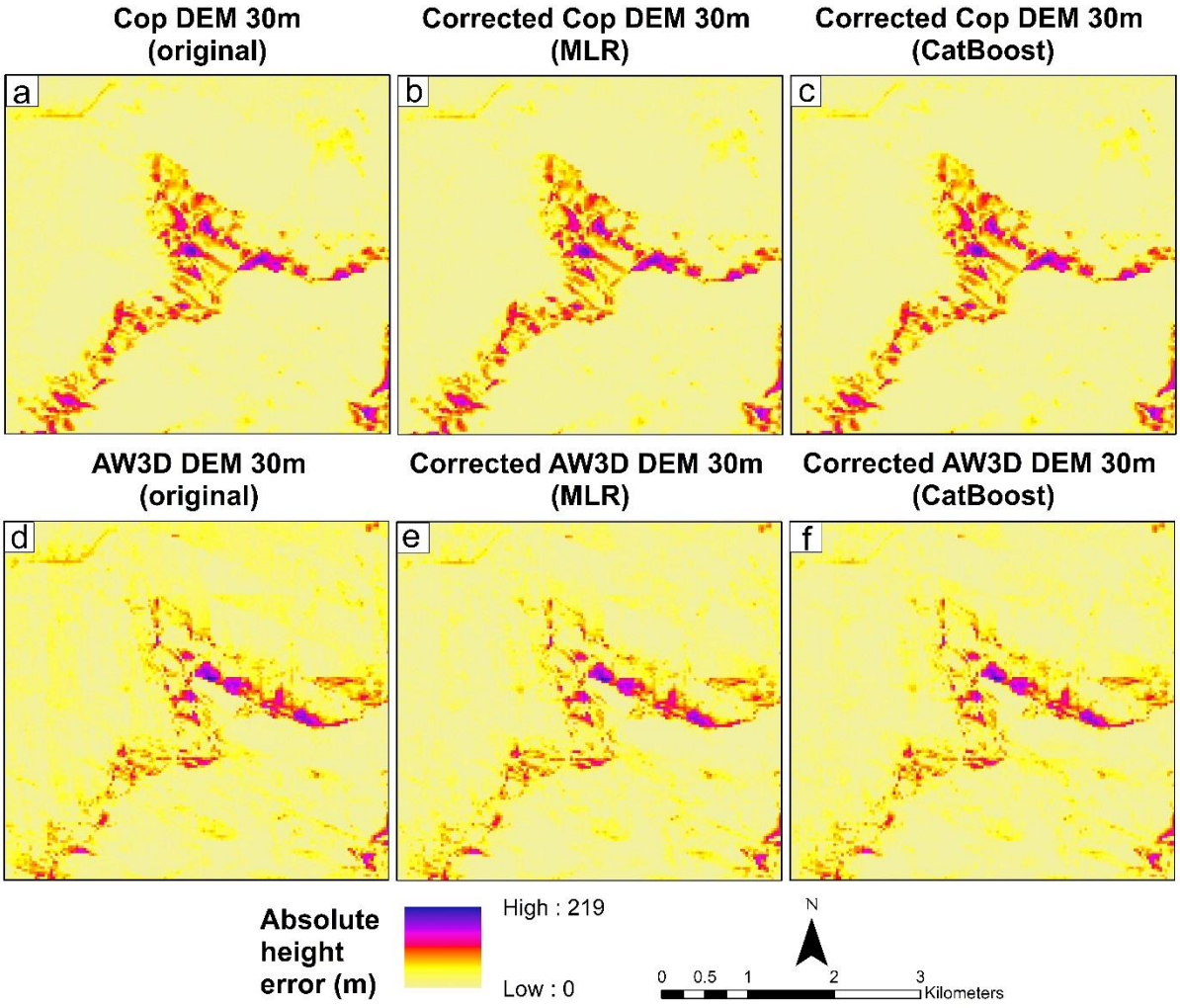


Figure 5.98: Absolute height error comparison of statistical-based and the best default GBDT-based Copernicus DEM correction at the 4th implementation site – mountain landscape

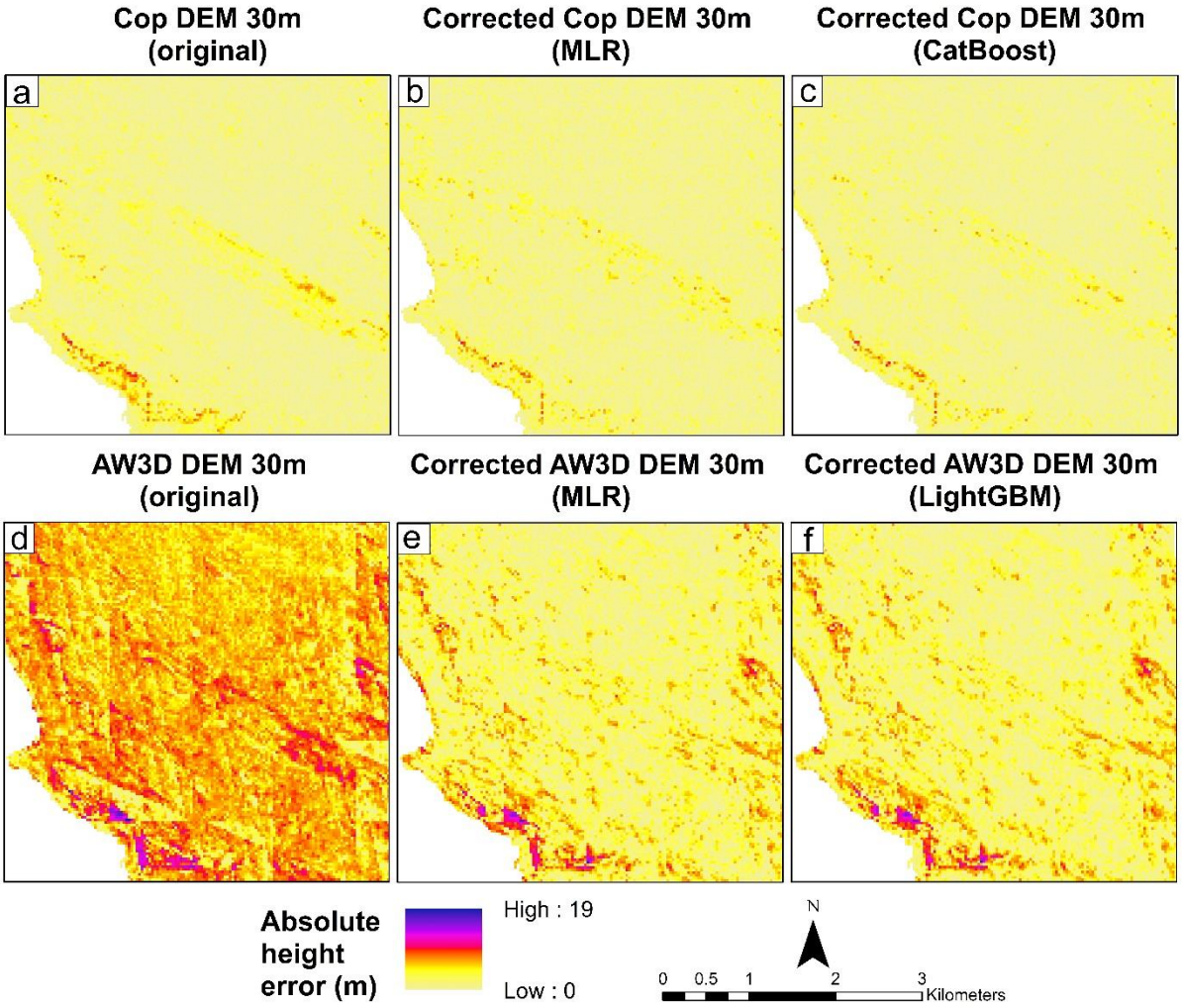


Figure 5.99: Absolute height error comparison of statistical-based and the best default GBDT-based Copernicus DEM correction at the 5th implementation site – peninsula landscape

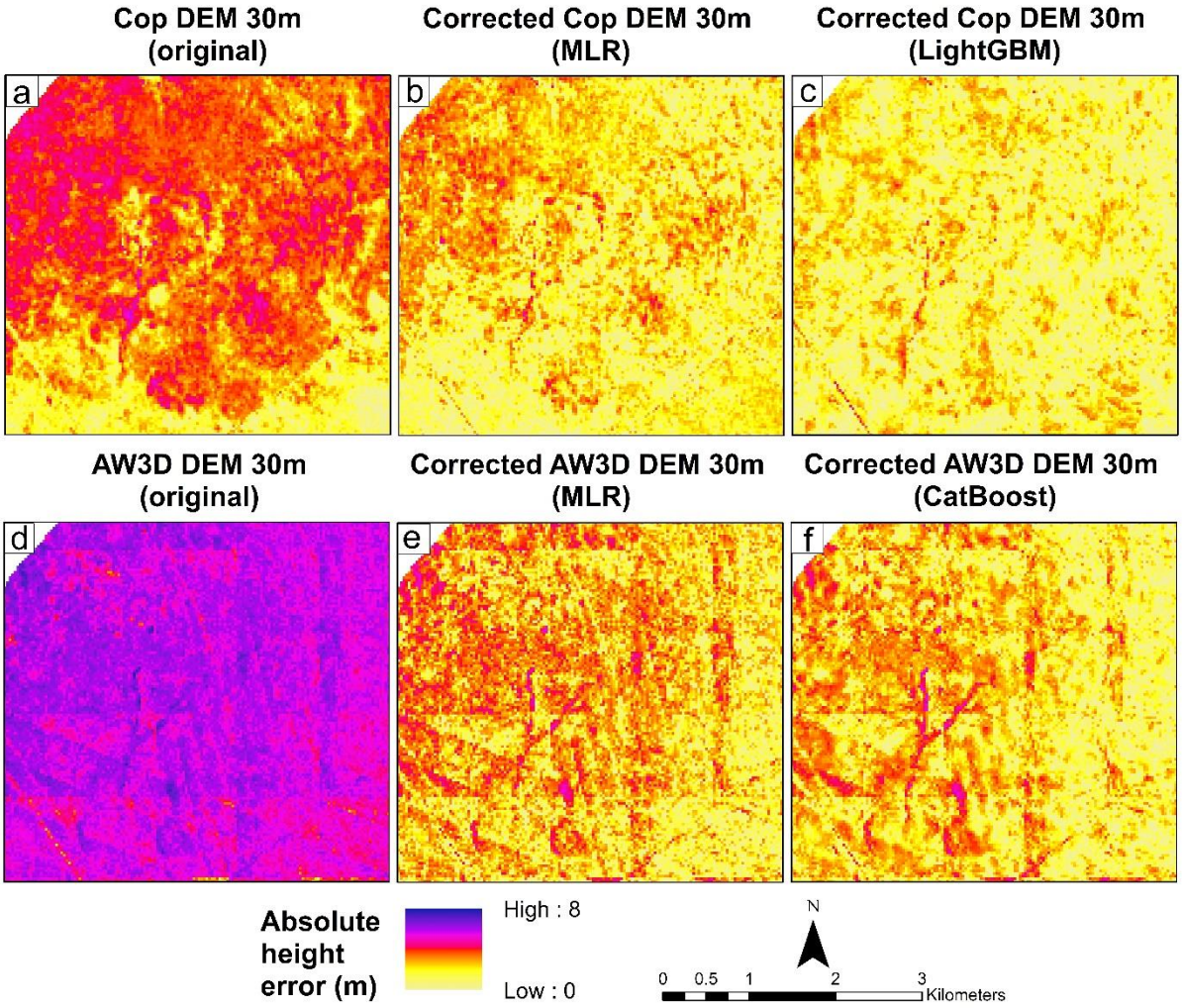


Figure 5.100: Absolute height error comparison of statistical-based and the best default GBDT-based Copernicus DEM correction at the 4th implementation site – grassland/ shrubland

5.5.4 Comparison with FABDEM and South Africa national DEM

In this section, the enhanced DEMs are compared with the internationally acclaimed FABDEM (Forest And Buildings removed Copernicus DEM), and the authoritative South African National DEM. At the 4th implementation site (Table 5.40), FABDEM and SA National DEM were superior to the original Copernicus DEM in the urban/industrial landscape. However, in the agricultural, mountain, peninsula and grassland/shrubland landscapes, FABDEM proved inferior to the original Copernicus DEM. SA National DEM also surpassed the accuracy of the original Copernicus DEM, the corrected Copernicus DEM and FABDEM at the urban/industrial and mountain landscapes, but proved inferior to the original and corrected Copernicus DEMs in the peninsula landscape.

In some areas, the Table Mountain is covered by trees and the urban areas have elevated objects. The South Africa National DEM has the advantage of a higher resolution (25 m), and its sensor is in better proximity with the terrain compared to the satellites. Thus, it has the potential to characterise the area better, and account for the trees and elevated objects better than the satellite DEMs. Moreover, InSAR suffers from the effects of foreshortening and layover in mountainous areas. and this could be a possible cause of the decreased accuracy in the Table Mountain landscape. Another issue is the SAR pulse resolution (footprint). In mountainous areas, especially those with a lot of local terrain variations, the pulse return may become a bit noisy. Photogrammetrically, a specific point is mapped on the terrain. This is normally done at a grid intersection plus at features of sharp terrain relief change (so breaklines may be included). The final 25 m DEM grid height will thus be determined from the surrounding 'spotshots', breaklines, and terrain high/low spots. While the corrected Copernicus DEM dominated in most landscapes, FABDEM did not outperform either the corrected Copernicus DEM or the SA National DEM. Although several wide-area or globally corrected DEMs have been produced by teams of researchers around the world, the sparseness of the data used in their refinement could limit their reliabilities for precise mapping within national, regional or local territories.

At the 5th implementation site (Table 5.41), the corrected Copernicus DEM dominated the assessment in the agricultural, peninsula and grassland/shrubland landscapes. Although the lowest RMSE of 7.229 m in the mountain landscape was achieved by SA National DEM, the corrected Copernicus DEM achieved the lowest MAE of 4.080 m. FABDEM dominated the urban/industrial area with the lowest MAE and RMSE of 0.480 m and 0.725 m respectively. According to Hawker et al. (2022), the corrections to FABDEM were applied in built-up areas and forested areas only, and they caution users to be mindful when using FABDEM in other environments, particularly in steep slopes. The developers of FABDEM (Hawker et al., 2022) also reported their use of a wide range of independent datasets over urban areas in training their correction model to improve its applicability. However, the differences in the achieved accuracies of the corrected Copernicus DEM and FABDEM are quite marginal. Figure 5.101 and Figure 5.102 showcase visualisations of RMSEs of the corrected Copernicus DEM compared with the globally corrected Copernicus DEM (FABDEM) and the South Africa National DEM at the 4th and 5th implementation sites respectively. Figure 5.103 - Figure 5.107 present absolute height error comparisons of the best GBDT-based Copernicus DEM correction versus FABDEM and the South Africa National DEM in different landscapes.

Table 5.40: Accuracy measures of the corrected Copernicus DEM compared with the globally corrected Copernicus DEM (FABDEM) and the South Africa National DEM at the 4th implementation site. The best accuracies are highlighted

Landscape	Original Copernicus DEM		Corrected Copernicus DEM (B.Cat)		FABDEM		SA National DEM	
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)
Urban/ industrial	1.389	1.952	0.608	0.941	0.736	0.999	**0.584**	**0.806**
Agricultural	0.456	1.466	**0.346**	**0.895**	0.840	1.595	0.798	1.074
Mountain	6.807	17.339	6.438	17.193	8.211	17.751	**5.469**	**10.018**
Peninsula	0.424	0.905	**0.352**	**0.693**	1.318	2.593	0.919	1.556
Grassland/ shrubland	1.357	1.520	**0.434**	**0.550**	1.384	1.557	0.721	0.900

Table 5.41: Accuracy measures of the corrected Copernicus DEM compared with the globally corrected Copernicus DEM (FABDEM) and the South Africa National DEM at the 5th implementation site. The best accuracies are highlighted

Landscape	Original Copernicus DEM		Corrected Copernicus DEM (B.Cat)		FABDEM		SA National DEM	
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)
Urban/ industrial	1.087	1.479	0.509	0.788	**0.480**	**0.725**	0.620	0.821
Agricultural land	0.420	1.230	**0.325**	**0.760**	0.886	1.486	0.839	1.164
Mountain	4.848	10.237	**4.080**	9.206	11.460	16.033	4.617	**7.229**
Peninsula	0.345	0.521	**0.292**	**0.430**	1.324	2.163	0.914	1.502
Grassland/ shrubland	0.497	0.563	**0.266**	**0.373**	0.677	0.877	0.675	0.930

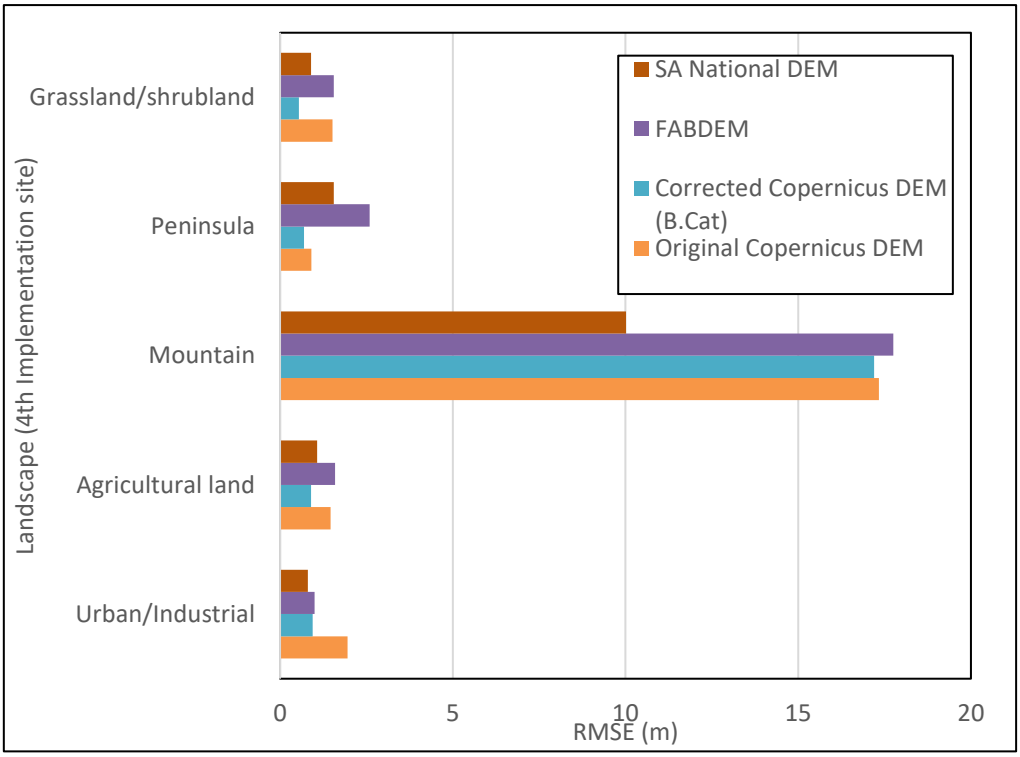


Figure 5.101: Visualisation of RMSEs of the corrected Copernicus DEM compared with the globally corrected Copernicus DEM (FABDEM) and the South Africa National DEM at the 4th implementation site

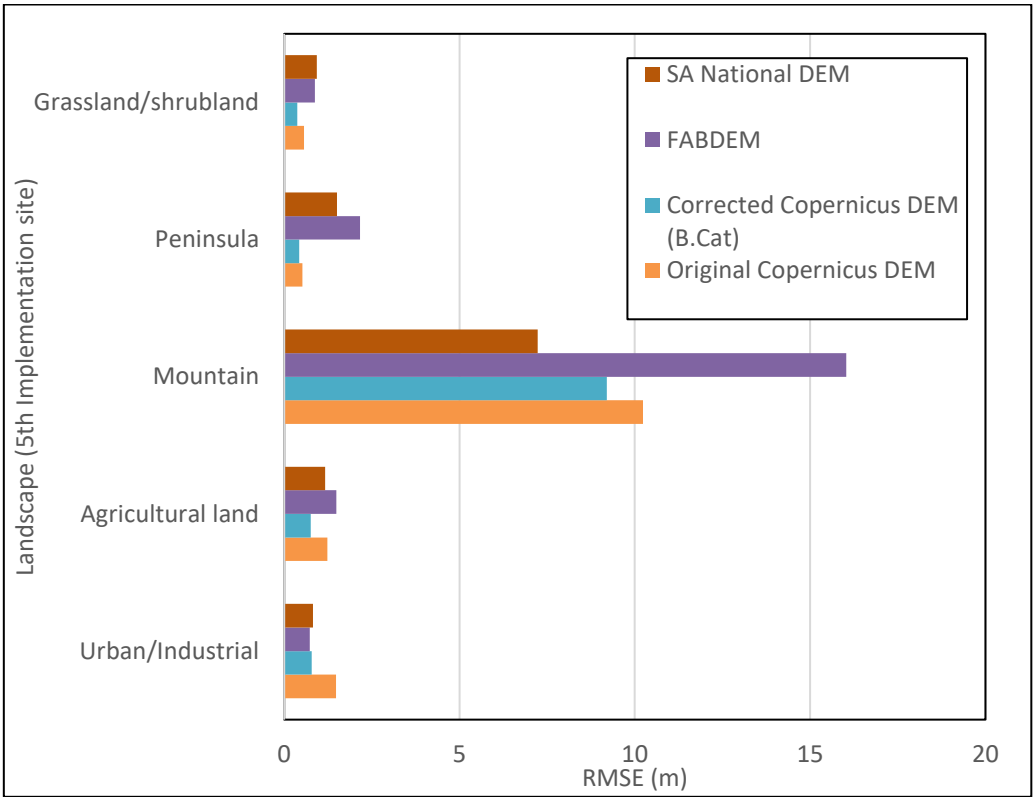


Figure 5.102: Visualisation of RMSEs of the corrected Copernicus DEM compared with the globally corrected Copernicus DEM (FABDEM) and the South Africa National DEM at the 5th implementation site

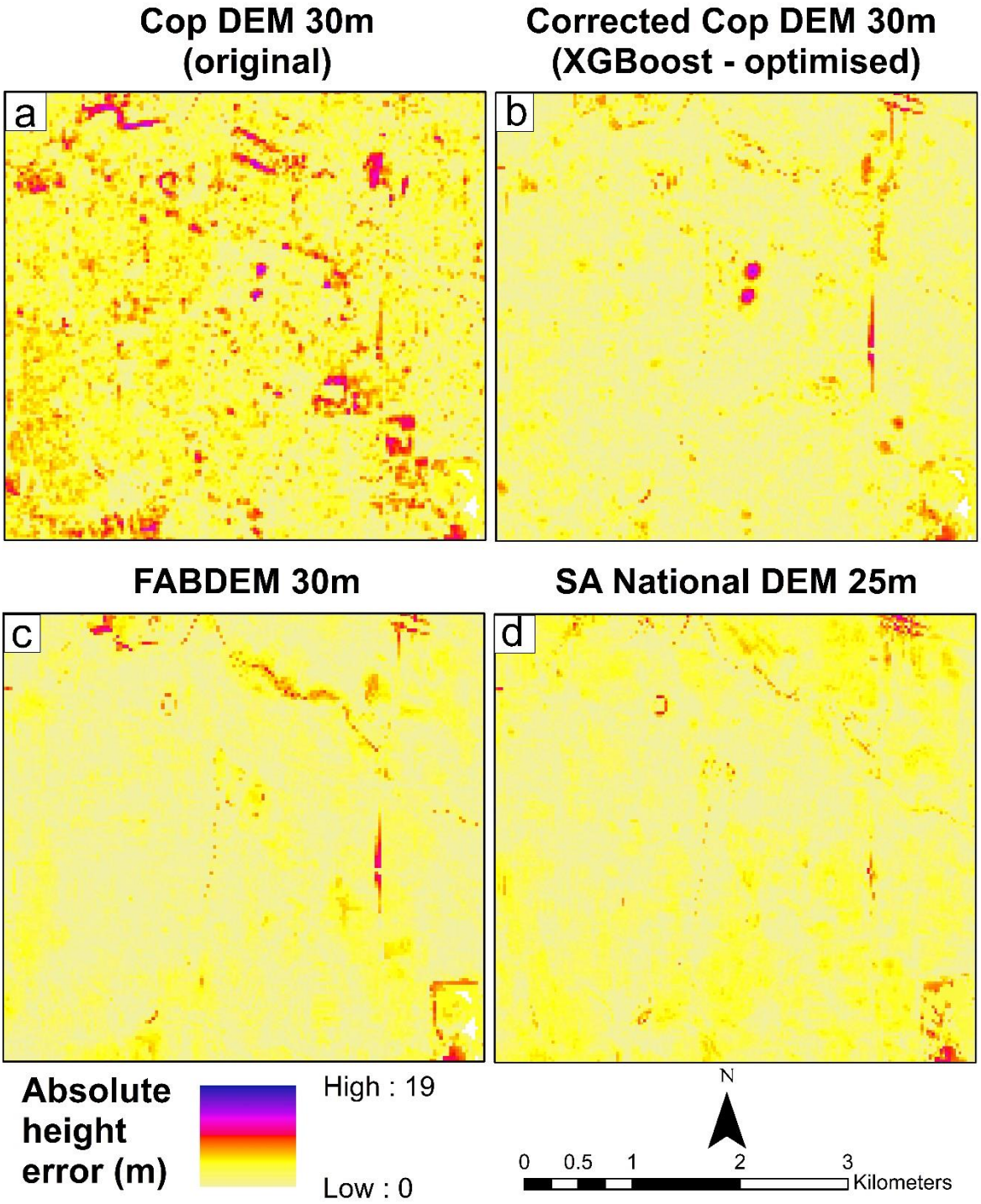


Figure 5.103: Absolute height error comparison of the best GBDT-based Copernicus DEM correction versus FABDEM and the South Africa National DEM at the 5th implementation site – urban/ industrial landscape

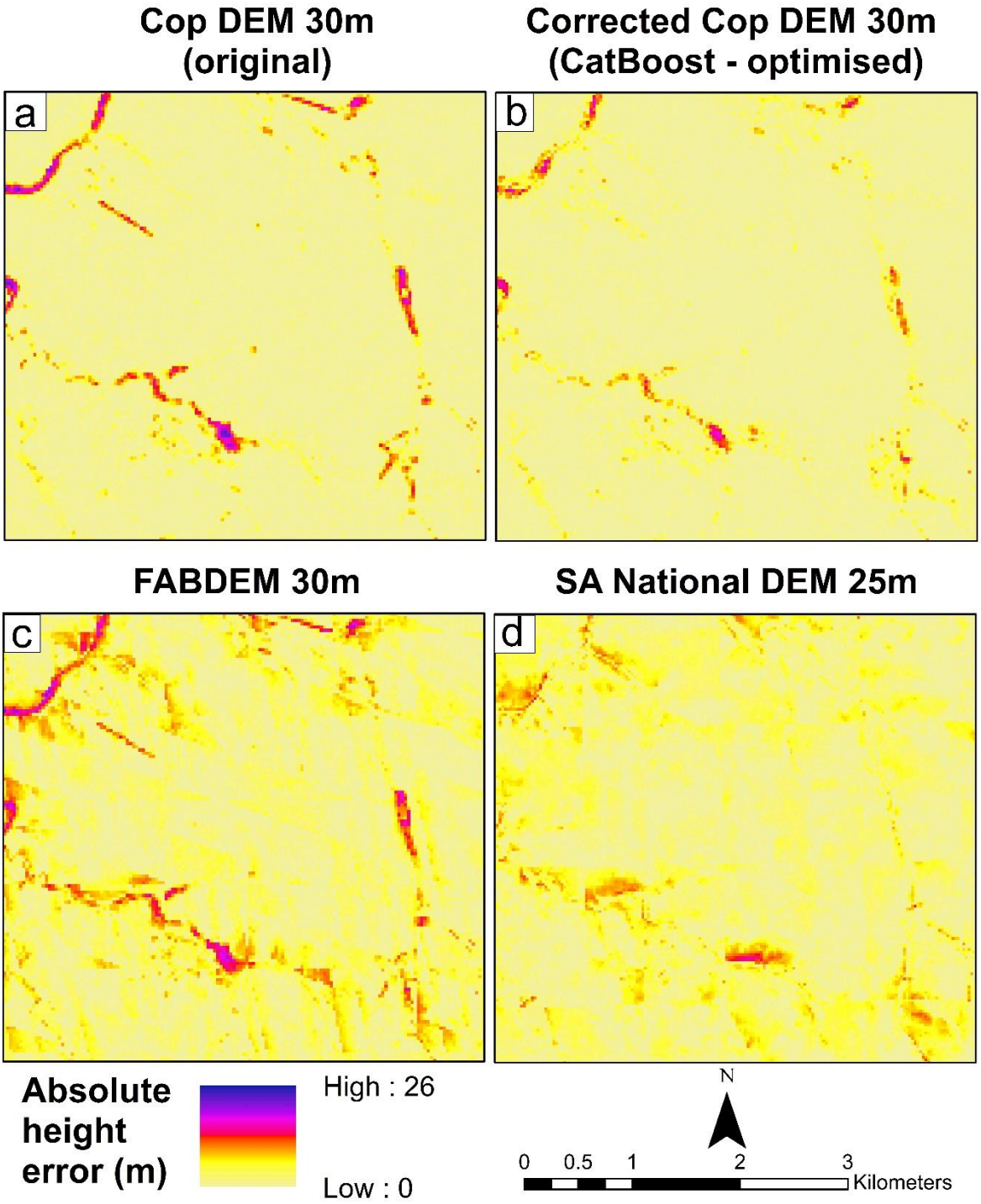


Figure 5.104: Absolute height error comparison of the best GBDT-based Copernicus DEM correction versus FABDEM and the South Africa National DEM at the 5th implementation site – agricultural land

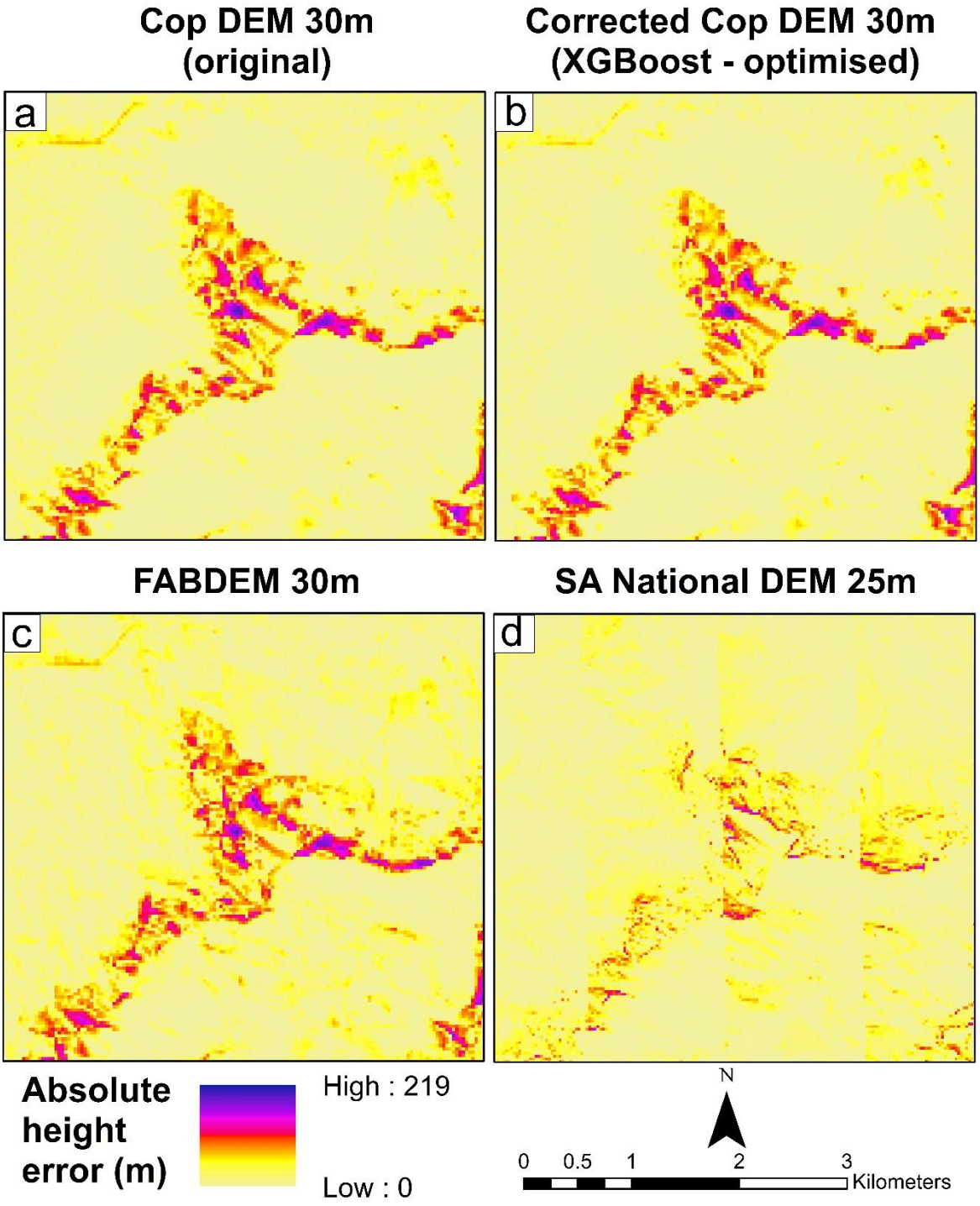


Figure 5.105: Absolute height error comparison of the best GBDT-based Copernicus DEM correction versus FABDEM and the South Africa National DEM at the 4th implementation site – mountain landscape

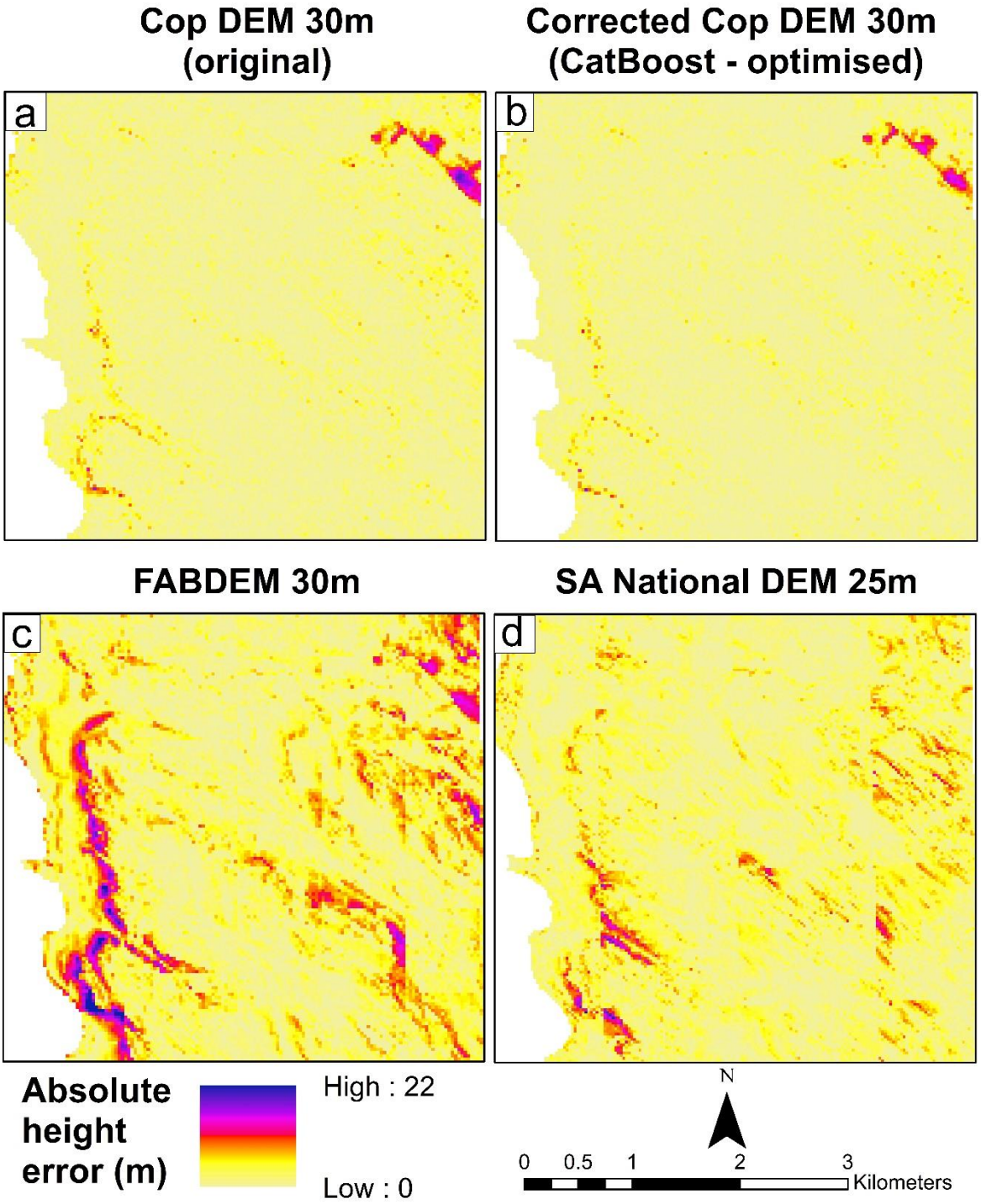


Figure 5.106: Absolute height error comparison of the best GBDT-based Copernicus DEM correction versus FABDEM and the South Africa National DEM at the 4th implementation site – peninsula landscape

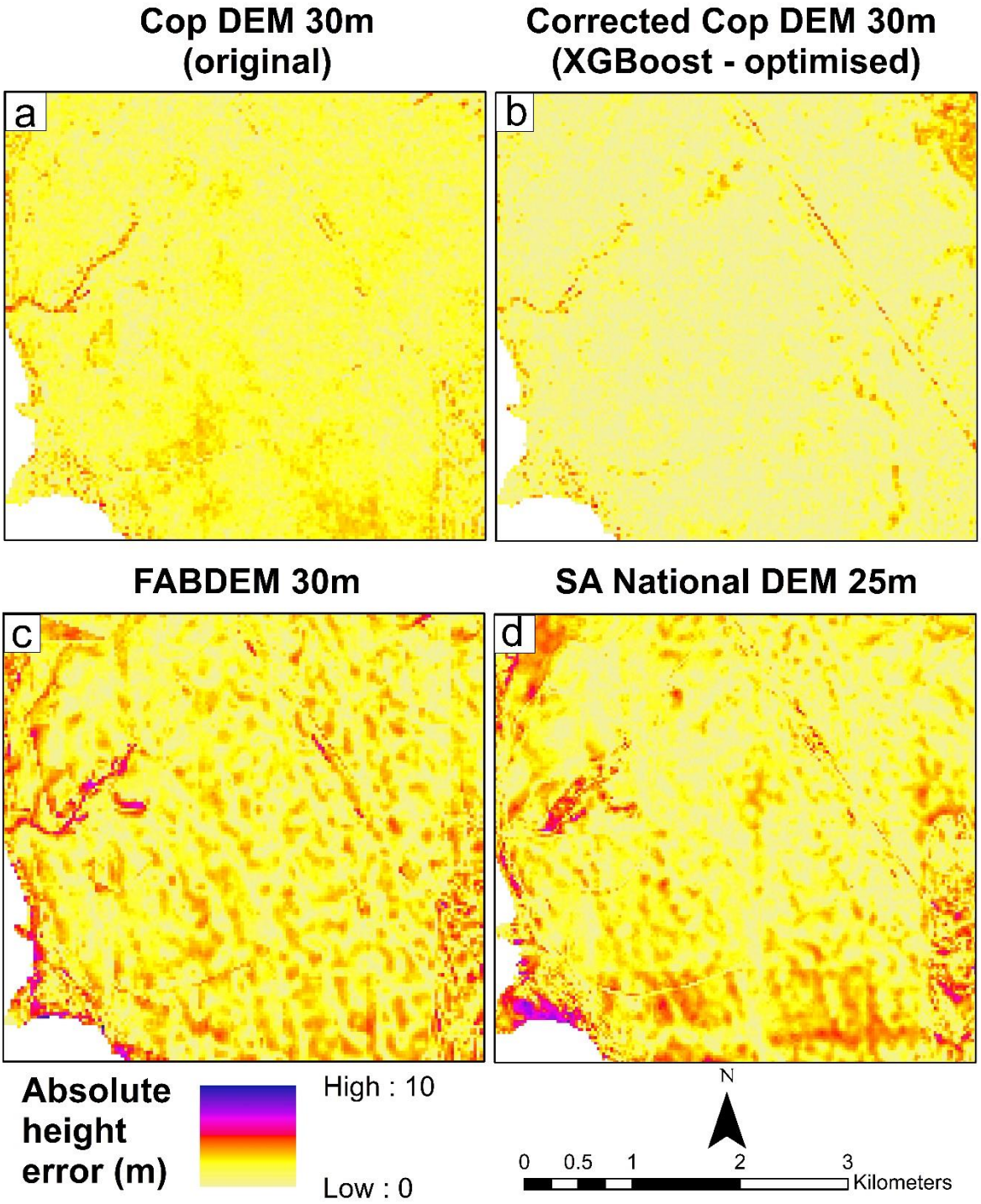


Figure 5.107: Absolute height error comparison of the best GBDT-based Copernicus DEM correction versus FABDEM and the South Africa National DEM at the 5th implementation site – grassland/shrubland

5.5.5 Comparison with previous studies

In this section, the achievements of the proposed DEM enhancement scheme are compared with results from previous studies that applied tree-based models for DEM correction. Most of the previous approaches adopted the bagging ensemble, random forest which has become

popular within the remote sensing community. However, the three boosting ensembles employed in the present study are more recent implementations of gradient boosting. The competitiveness of the proposed DEM enhancement scheme is evident in Table 5.42.

Table 5.42: Comparison of proposed DEM enhancement scheme with previous tree-based machine learning approaches

Study	Country of study	Tree-based model	Percentage improvement/error reduction
The present study	South Africa	Gradient boosted decision tree-based (XGBoost, LightGBM, CatBoost) feature-level fusion framework.	**AW3D DEM:** 14.6 – 81.6% MAE reduction 5.5 – 77.7% RMSE reduction **Copernicus DEM:** 5.2 – 68% MAE reduction 0.7 – 63.8% RMSE reduction
Meadows & Wilson (2021)	New Zealand	Random Forest	55% RMSE reduction
Chen et al. (2020)	China	Random Forest	8.4% RMSE reduction
Hu & Ji (2022)	China	Random Forest	**AW3D:** 14.67% RMSE reduction **ASTER:** 10.58% RMSE reduction
Hawker et al. (2022)	Global	Random Forest	**Copernicus DEM:** 44.1% MAE reduction in forests 30.4% MAE reduction in built-up areas
Liu et al. (2021)	China, Germany, United Kingdom	Random Forest	15–67%, and 57 – 76% error reduction, in two separate experiments

5.5.6 Discussion of comparative assessment results

The South Africa National DEM has a slightly higher spatial resolution of 25 m and is photogrammetrically sourced or produced from vectorised contours. Because of its higher resolution, it is expected to characterise the terrain to a greater level of detail. Moreover, the data sources for the National DEM are expected to offer better accuracy than satellite-derived datasets. However, Copernicus GLO-30 is a satellite-derived 30m DEM that originated from InSAR technology. The comparisons show that in most landscapes, the enhancement scheme outperformed both FABDEM and the South Africa national DEM. However, the 25 m South

Africa national DEM outperformed the enhanced DEMs in the two mountainous landscapes and at one of the urban sites. Interferometric DEMs are subject to some errors in mountainous landscapes (foreshortening and layover). The impact on the DEM accuracy is dependent upon the sensor-scene geometry, but this could be a possible cause of the decreased accuracy in mountain landscapes. Another issue is the SAR pulse resolution (footprint). In mountains, especially those with a lot of local terrain variations, the pulse return may become a bit noisy. Photogrammetrically, a specific point is mapped on the terrain. This is normally done at a grid intersection plus at features of sharp terrain relief change (so breaklines may be included). The final 25 m DEM grid elevation was likely determined from the surrounding 'spotshots', breaklines, and terrain high/low spots.

5.6 Chapter Summary

Through a series of systematic evaluations, Copernicus DEM was selected among four global DEMs for enhancement using the presented framework. Summarily, the tree-based ensembles delivered competitive and complimentary results, and the enhancement framework allows flexibility depending on the user requirements, specifications and desired end results.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

For decades, the issue of positional errors in satellite remote sensing data has confounded scientists and researchers; it remains a recurring theme in the remote sensing community all around the world. The presence of vertical errors in global DEMs has been a challenge and a source of concern to end-users of satellite remote sensing products, geomatics practitioners and the remote sensing community at large. It compromises the utility of global DEMs for diverse applications in the national, regional and global domains. The aim of this research was the enhancement of satellite-derived multi-sensor (interferometric and photogrammetric) global DEMs through the implementation of a tree-based ensemble feature-level fusion scheme that exploits the interactions between terrain and land cover parameters. This was achieved through a series of objectives that ultimately addressed the research questions. Firstly, a quality assessment was carried out for two interferometric (NASADEM and Copernicus) and two photogrammetric (ASTER and AW3D) global DEMs to inform their selection for further analysis. Subsequently, a data fusion methodology for enhancement of Copernicus and AW3D were developed using tree-based ensemble (and GBDT-based) feature-level fusion with various configurations. Finally, the predicted outcomes were interpreted using machine learning explainability measures, and performance assessments were conducted on the DEM enhancement scheme to prove its efficacy.

The backbone of the framework is a gradient boosted decision tree (GBDT)-based feature-level fusion of terrain and land cover parameters that incorporates Bayesian optimisation for DEM error prediction and correction. Training data for the algorithms are drawn from a wide array of landscapes in Cape Town, South Africa including urban/industrial, agricultural, mountain, peninsula and grassland/shrubland landscapes. These sites are also representative of other global landscapes where the issue of global DEM correction remains an urgent priority.

The following conclusions are drawn from this research:

1. The DEM enhancement scheme was successful in the reduction of vertical errors in both Copernicus and AW3D DEMs through the application of point-by-point, location-specific corrections.

2. Two scenarios were considered in the implementation of gradient boosting: the use of default hyperparameters versus Bayesian-optimised hyperparameters. Both scenarios proved effective in significantly reducing the overall elevation bias of the original DEMs.
3. The accuracies achieved by XGBoost, LightGBM and CatBoost in the GBDT-based DEM enhancement framework were comparable in several landscapes. The ultimate choice of which model to adopt could depend on the computational resources available to users. For example, LightGBM proved to be several times faster than both XGBoost and CatBoost in the training and Bayesian optimisation. However, in most cases, CatBoost outperformed both XGBoost and LightGBM in prediction accuracy.
4. The explainability analysis shows that less popular terrain parameters such as VRM and TRI could be very influential for the prediction of DEM error.
5. Comparative assessments showed that the proposed method outperformed the DEM error diminution capability of several techniques that had been proposed in previous studies. As a proof of concept, the robustness and efficacy of the proposed enhancement scheme were proven at several implementation sites. The proposed enhancement scheme led to significant and competitive improvements in DEM accuracy, as high as eighty percent reduction in the elevation bias. The findings also provide insights into the cause-and-effect relationship between terrain and land cover parameters that are known influencers of DEM error.
6. The enhancement scheme offers a high level of flexibility in the input parameters, can be applied to diverse landscapes around the world with similar terrain characteristics as the training sites, and for predicting DEM error even on landscapes where ground truth data is not available. Moreover, the proposed DEM enhancement scheme is scalable and can be adopted by international space and mapping agencies for improving the vertical accuracies of future global DEM versions and releases to improve their accuracies. It offers a low-cost option to low and middle income (LMIC) countries to make more informed adoption of global DEMs for revising their nationwide topographic maps and for other related applications.
7. The DEM enhancement scheme developed in this research has the capability to expand the applicability of global DEMs by eliminating severe elevation biases and cutting down the RMSE by significant proportions in several landscapes as shown in the qualitative and quantitative assessments.

8. The applicability of the enhancement scheme for 3D terrain analysis and visualisation is proven, with the implication that it can serve as a potential alternative or substitute for costly airborne topographic surveys in the revision of topographic maps by national mapping agencies in data-scarce regions of Africa and the global south.
9. Supervised machine learning regression is a powerful and effective approach for modelling complex and non-linear terrain parameters even in challenging, inaccessible and difficult landscapes. Tree-based ensemble machine learning and recent implementations of gradient boosting are very powerful for reducing the uncertainty in digital elevation datasets. Given the sheer amount of environmental, hydrological and geological applications that rely on free global DEMs, DEM correction with machine learning will remain a strategic research mandate in the remote sensing community.
10. It is anticipated that this pioneering work will influence further applications of tree-based ensembles in remote sensing data fusion and DEM enhancement.

6.2 Research Limitations

Training data for the enhancement scheme were drawn from diverse landscapes of Cape Town. Although the training data are topographically diverse, some caution must be applied in the transferability and deployment of the models elsewhere, for example in heavily forested and glacial environments. Also, there is no one-size-fits-all solution for model optimisation and there is always room for improvement.

6.3 Recommendations

The following recommendations are made for future research:

1. The proposed enhancement scheme can be adopted by remote sensing research consortia for producing the next globally corrected DEM products, from the Copernicus DEM, or other DEMs with global coverage. Ground-level and canopy height metrics from space-borne LiDAR (e.g., the Ice, Cloud and Land Elevation Satellite – ICESat-2; and the Global Ecosystem Dynamics Investigation – GEDI) are vital for generating target layers for such training (i.e. height error maps) at a global scale.
2. While high resolution DEMs or DTMs are available in many advanced countries, low and middle income (LMICs) continue to struggle with funding constraints to pursue such high-resolution mapping. Thus, the proposed enhancement scheme is a low-cost and viable alternative for national mapping agencies in LMICs to improve the vertical

accuracy of readily available global DEMs for use in their national geospatial infrastructure and mapping projects.

3. The enhancement of open-access satellite-derived DEMs is suggested as an alternative to national DEMs which may be inaccessible to interested stakeholders and end-users.
4. The results also indicate that ML algorithms are likely to be biased towards the terrain characteristics of the training data fed into them. Thus, further research is required to understand the topographic and geomorphological complexity in different landscapes and how this affects the generalisability of ML approaches for DEM correction over wide areas.
5. Despite the computational burden of hyperparameter optimisation, it is recommended for achieving the most optimal results.
6. The developed methodology is not only applicable to global DEMs, but is also recommended for improving the accuracy of national DEMs and even airborne LiDAR.
7. Finally, recent progress in space exploration and topographic mapping of planetary bodies has created opportunities for the application of the DEM enhancement methodologies for enhanced terrain analysis of planetary bodies.

REFERENCES

Abolfazli, A., Brechmann, A., Wolff, S., & Spiliopoulou, M. (2020). Machine learning identifies the dynamics and influencing factors in an auditory category learning experiment. *Scientific Reports*, 10(1). https://doi.org/10.1038/S41598-020-61703-X

Abrams, M., Crippen, R., & Fujisada, H. (2020). ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). *Remote Sensing*, 12(7), 1156. https://doi.org/10.3390/RS12071156

Acharya, T. D., Yang, I. T., & Lee, D. H. (2018). Comparative Analysis of Digital Elevation Models between AW3D30, SRTM30 and Airborne LiDAR: A Case of Chuncheon, South Korea. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 36(1), 17–24. https://doi.org/10.7848/KSGPC.2018.36.1.17

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Adler, A. I., & Painsky, A. (2022). Feature Importance in Gradient Boosting Trees with Cross-Validation Feature Selection. *Entropy*, 24(5), 687. https://doi.org/10.3390/E24050687

Afana, A. (2011). *Delineation of channel networks from Digital Elevation Models (DEMs)* [University of Almeria]. http://www.eeza.csic.es/documentos/tesis/Tesis doctoral (Afana, 2011).pdf

Agrawal, R., Das, A., & Rajawat, A. S. (2018). Accuracy Assessment of Digital Elevation Model Generated by SAR Stereoscopic Technique Using COSMO-Skymed Data. *Journal of the Indian Society of Remote Sensing*, 46(10), 1739–1747. https://doi.org/10.1007/S12524-018-0835-6/FIGURES/10

Ågren, A. M., Larson, J., Paul, S. S., Laudon, H., & Lidberg, W. (2021). Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape. *Geoderma*, 404, 115280. https://doi.org/10.1016/J.GEODERMA.2021.115280

Ahmadi, H., & Pekkan, E. (2021). Fault-Based Geological Lineaments Extraction Using Remote Sensing and GIS—A Review. *Geosciences*, 11(5), 183. https://doi.org/10.3390/GEOSCIENCES11050183

Airbus. (2020a). *Copernicus DEM Copernicus Digital Elevation Model Product Handbook*.

Airbus. (2020b). *Copernicus DEM Copernicus Digital Elevation Model Validation Report*.

Aksoy, N., & Genc, I. (2023). Predictive models development using gradient boosting based methods for solar power plants. *Journal of Computational Science*, 67, 101958. https://doi.org/10.1016/J.JOCS.2023.101958

Al-Khafaji, M., Saeed, F. H., & Al-Ansari, N. (2020). The Interactive Impact of Land Cover and DEM Resolution on the Accuracy of Computed Streamflow Using the SWAT Model. *Water, Air, and Soil Pollution*, 231(8), 1–17. https://doi.org/10.1007/S11270-020-04770-0/TABLES/4

Alademomi, A. S., Okolie, C. J., Daramola, O. E., Akinnusi, S. A., Adediran, E., Olanrewaju, H. O., Alabi, A. O., Salami, T. J., & Odumosu, J. (2022). The interrelationship between

LST, NDVI, NDBI, and land cover change in a section of Lagos metropolis, Nigeria. *Applied Geomatics*, 14(2), 299–314. https://doi.org/10.1007/S12518-022-00434-2/FIGURES/9

Alam, M. (2021). *A guide to XGBoost hyperparameters*. Towards Data Science. https://towardsdatascience.com/a-guide-to-xgboost-hyperparameters-87980c7f44a9#:~:text=n_estimators %3A specifies the number of,ratio of the training sample

Albert, G., & Ammar, S. (2021). Application of random forest classification and remotely sensed data in geological mapping on the Jebel Meloussi area (Tunisia). *Arabian Journal of Geosciences*, 14(21), 1–13. https://doi.org/10.1007/S12517-021-08509-X/TABLES/6

Anjum, M., Khan, K., Ahmad, W., Ahmad, A., Amin, M. N., & Nafees, A. (2022). New SHapley Additive ExPlanations (SHAP) Approach to Evaluate the Raw Materials Interactions of Steel-Fiber-Reinforced Concrete. *Materials*, 15(18), 6261. https://doi.org/10.3390/MA15186261

Anuradha, & Gupta, G. (2014). A self explanatory review of decision tree classifiers. *International Conference on Recent Advances and Innovations in Engineering, ICRAIE*. https://doi.org/10.1109/ICRAIE.2014.6909245

Aronoff, S. (1989). Geographic information systems: A management perspective. *Geocarto International*, 4(4), 58–58. https://doi.org/10.1080/10106048909354237

ASPRS. (2014). *ASPRS Positional Accuracy Standards for Digital Geospatial Data*. https://www.asprs.org/a/society/committees/standards/Positional_Accuracy_Standards.pdf

ASPRS LiDAR Committee (PAD): (2004). *ASPRS. Vertical Accuracy Reporting for LiDAR Data*.

Atit, K., Weisberg, S. M., Newcombe, N. S., & Shipley, T. F. (2016). Learning to interpret topographic maps: Understanding layered spatial information. *Cognitive Research: Principles and Implications*, 1(1), 1–18. https://doi.org/10.1186/S41235-016-0002-Y/FIGURES/8

Audenino, P., Rognant, L., Chassery, J. M., & Planes, J. G. (2001). Fusion strategies for high resolution urban DEM. *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, DFUA 2001*, 90–94. https://doi.org/10.1109/DFUA.2001.985733

Azimi, J., Jalali, A., & Fern, X. Z. (2012). Hybrid Batch Bayesian Optimization. *Proceedings of the 29th International Conference on Machine Learning, ICML*, 2, 1215–1222. https://arxiv.org/abs/1202.5597v3

Bagheri, H., Schmitt, M., & Zhu, X. (2017). Fusion of TanDEM-X and Cartosat-1 DEMS using TV-norm regularization and ANN-predicted weights. *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 3369–3372. https://doi.org/10.1109/IGARSS.2017.8127720

Bagheri, H., Schmitt, M., & Zhu, X. X. (2018a). Fusion of TanDEM-X and Cartosat-1 elevation data supported by neural network-predicted weight maps. *Isprs Journal of Photogrammetry and Remote Sensing*, 144(August), 285–297.

https://doi.org/10.1016/j.isprsjprs.2018.07.007

Bagheri, H., Schmitt, M., & Zhu, X. X. (2018b). Urban TanDEM-X Raw DEM Fusion Based ON TV-L1 and Huber Models. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018-July(June), 7251–7254. https://doi.org/10.1109/IGARSS.2018.8518870

Bahmani, M. . (2023). *Understanding LightGBM Parameters (and How to Tune Them)*. MLOps Blog. https://neptune.ai/blog/lightgbm-parameters-guide

Bai, X., Liu, C., Ren, P., Zhou, J., Zhao, H., & Su, Y. (2015). Object classification via feature fusion based marginalized kernels. *IEEE Geoscience and Remote Sensing Letters*, 12(1), 8–12. https://doi.org/10.1109/LGRS.2014.2322953

Bakhareva, N., Shukhman, A., Matveev, A., Polezhaev, P., Ushakov, Y., & Legashev, L. (2019). Attack Detection in Enterprise Networks by Machine Learning Methods. *Proceedings - International Russian Automation Conference, RusAutoCon*. https://doi.org/10.1109/RUSAUTOCON.2019.8867696

Barbarella, M., Di Benedetto, A., & Fiani, M. (2022). A Method for Obtaining a DEM with Curved Abscissa from MLS Data for Linear Infrastructure Survey Design. *Remote Sensing*, 14(4), 889. https://doi.org/10.3390/RS14040889

Barber, C. P., & Shortridge, A. (2013). Lidar Elevation Data for Surface Hydrologic Modeling: Resolution and Representation Issues. *Cartography and Geographic Information Science*, 32(4), 401–410. https://doi.org/10.1559/152304005775194692

Barrow, D. K., & Crone, S. F. (2016). A comparison of AdaBoost algorithms for time series forecast combination. *International Journal of Forecasting*, 32(4), 1103–1119. https://doi.org/10.1016/J.IJFORECAST.2016.01.006

Bauer, E., & Kohavi, R. (1999). Empirical comparison of voting classification algorithms: bagging, boosting, and variants. *Machine Learning*, 36(1), 105–139. https://doi.org/10.1023/A:1007515423169/METRICS

Baugh, C., Bates, P., Schumann, G., & Trigg, M. (2013). SRTM vegetation removal and hydrodynamic modeling accuracy. *Water Resources Research*, 49, 5276–5289. https://doi.org/10.1002/WRCR.20412

Beg, A. A. F. (2020). Spatial analysis of the geomorphic evolution of Tigris River basin using developed ArcGIS-Morphometric toolbox. *IOP Conference Series: Materials Science and Engineering*, 737(1), 12235. https://doi.org/10.1088/1757-899X/737/1/012235

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. https://doi.org/10.1016/J.ISPRSJPRS.2016.01.011

Benahmed Daho, S. A. (2009). Evaluation of the Earth Gravity Model EGM2008 in Algeria. *Newon's Bulletin*, *4*, 172–184.

Benediktsson, J. A. (1997). Hybrid consensus theoretic classification. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4), 833–843. https://doi.org/10.1109/36.602526

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for Hyper-Parameter

Optimization. *Advances in Neural Information Processing Systems*, 24.

Bettiol, G. M., Ferreira, M. E., Motta, L. P., Cremon, É. H., & Sano, E. E. (2021). Conformity of the NASADEM_HGT and ALOS AW3D30 DEM with the Altitude from the Brazilian Geodetic Reference Stations: A Case Study from Brazilian Cerrado. *Sensors*, 21(9), 2935. https://doi.org/10.3390/S21092935

Bhardwaj, A., Jain, K., & Chatterjee, R. S. (2019). Generation of high-quality digital elevation models by assimilation of remote sensing-based DEMs. *Journal of Applied Remote Sensing*, 13(04), 44502. https://doi.org/10.1117/1.JRS.13.4.044502

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227. https://doi.org/10.1007/S11749-016-0481-7/FIGURES/4

Bogale, A. (2021). Morphometric analysis of a drainage basin using geographical information system in Gilgel Abay watershed, Lake Tana Basin, upper Blue Nile Basin, Ethiopia. *Applied Water Science*, 11(7), 1–7. https://doi.org/10.1007/S13201-021-01447-9/TABLES/2

Borisov, V., Leemann, T., Sessler, K., Haug, J., Pawelczyk, M., Kasneci, G., Seßler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2022). Deep Neural Networks and Tabular Data: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*. https://doi.org/10.1109/TNNLS.2022.3229161

Brassington, G. (2017). Mean absolute error and root mean square error: which is the better metric for assessing model performance? *Geophysical Research Abstracts*, 19, 2017–3574.

Braun, A. (2021). Retrieval of digital elevation models from Sentinel-1 radar data - Open applications, techniques, and limitations. *Open Geosciences*, 13(1), 532–569. https://doi.org/10.1515/GEO-2020-0246/ASSET/GRAPHIC/J_GEO-2020-0246_FIG_009.JPG

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. https://doi.org/10.1007/BF00058655/METRICS

Breiman, L. (1999). Pasting small votes for classification in large databases and on-line. *Machine Learning*, 36(1), 85–103. https://doi.org/10.1023/A:1007563306331/METRICS

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. *Classification and Regression Trees*, 1–358. https://doi.org/10.1201/9781315139470/CLASSIFICATION-REGRESSION-TREES-LEO-BREIMAN

Brownlee, J. (2016). *How to Tune the Number and Size of Decision Trees with XGBoost in Python*. Machine Learning Mastery. https://machinelearningmastery.com/tune-number-size-decision-trees-xgboost-python/

Buckley, S. M., Agram, P. S., Belz, J. E., Crippen, R. E., Gurrola, E. M., Hensley, S., Kobrick, M., Lavalle, M., Martin, J. M., Neumann, M., Nguyen, Q. D., Rosen, P. A., Shimada, J. G., Simard, M., & Tung, W. W. (2020). NASADEM: User Guide. *Nasa/Jpl*, *January*, 48.

https://lpdaac.usgs.gov/documents/592/NASADEM_User_Guide_V1.pdf

Buckley, S. M., Agram, P. S., Belz, J. E., Crippen, R. E., Gurrola, E. M., Hensley, S., Kobrick, M., Lavalle, M., Martin, J. M., Neumann, M., Nguyen, Q. D., Rosen, P. A., Shimada, J. G., Simard, M., & Tung, W. W. (2022). *NASADEM*.

Bühlmann, P. (2012). Bagging, Boosting and Ensemble Methods. *Handbook of Computational Statistics*, 985–1022. https://doi.org/10.1007/978-3-642-21551-3_33

Burrough, P., & McDonnell, R. (1998). *Principles of geographical information systems*. https://semanticscholar.org/paper/d8e0b6b225b36cac23608b41a51e13ddb2746cbd

Canadian Soil Information Service. (2013). *Slope Gradient*. https://sis.agr.gc.ca/cansis/nsdb/slc/v3.2/cmp/slope.html

Capaldo, P., Crespi, M., Fratarcangeli, F., Nascetti, A., & Pieralice, F. (2011). High-Resolution SAR Radargrammetry: A First Application With COSMO-SkyMed SpotLight Imagery. *IEEE Geoscience and Remote Sensing Letters*, 8, 1100–1104. https://doi.org/10.1109/LGRS.2011.2157803

Carrera-Hernández, J. J. (2021). Not all DEMs are equal: An evaluation of six globally available 30 m resolution DEMs with geodetic benchmarks and LiDAR in Mexico. *Remote Sensing of Environment*, 261, 112474. https://doi.org/10.1016/J.RSE.2021.112474

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. *ACM International Conference Proceeding Series*, 148, 161–168. https://doi.org/10.1145/1143844.1143865

CatBoost. (2023). *Overview*. CatBoost. https://catboost.ai/en/docs/references/training-parameters/

CD: NGI. (2019). *Standard for Digital Elevation Model Data*.

Cenci, L., Galli, M., Palumbo, G., Sapia, L., Santella, C., & Albinet, C. (2021). DESCRIBING THE QUALITY ASSESSMENT WORKFLOW DESIGNED FOR DEM PRODUCTS DISTRIBUTED VIA THE COPERNICUS PROGRAMME. CASE STUDY: THE ABSOLUTE VERTICAL ACCURACY OF THE COPERNICUS DEM DATASET IN SPAIN. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 6143–6146. https://doi.org/10.1109/IGARSS47720.2021.9554393

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. https://doi.org/10.5194/GMD-7-1247-2014

Chandler, G., & Merry, C. L. (2010). The South African Geoid 2010: SAGEOID10. *PositionIT*. www.positionIT.co.za

Chang, K. T., & Tsai, B. W. (2013). The Effect of DEM Resolution on Slope and Aspect Mapping. *Cartography and Geographic Information Systems*, 18(1), 69–77. https://doi.org/10.1559/152304091783805626

Chen, C., Yang, S., & Li, Y. (2020). Accuracy Assessment and Correction of SRTM DEM Using ICESat/GLAS Data under Data Coregistration. *Remote Sensing*, 12(20), 3435.

https://doi.org/10.3390/RS12203435

Chen, J., Luo, D. L., & Mu, F. X. (2009). An improved ID3 decision tree algorithm. *Proceedings of 2009 4th International Conference on Computer Science and Education, ICCSE*, 127–130. https://doi.org/10.1109/ICCSE.2009.5228509

Chen, P., Wang, B., Wu, Y., Wang, Q., Huang, Z., & Wang, C. (2023). Urban river water quality monitoring based on self-optimizing machine learning method using multi-source remote sensing data. *Ecological Indicators*, 146, 109750. https://doi.org/10.1016/J.ECOLIND.2022.109750

Chen, S.-Z., Feng, D.-C., Wang, W.-J., & Taciroglu, E. (2022). Probabilistic Machine-Learning Methods for Performance Prediction of Structure and Infrastructures through Natural Gradient Boosting. *Journal of Structural Engineering*, 148(8), 4022096. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003401

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13-17-August-2016, 785–794. https://doi.org/10.1145/2939672.2939785

Chesworth, W., Arbestain, M. ., Macías, F., Spaargaren, O., Spaargaren, O., Mualem, Y., Morel-Seytoux, H. ., Horwath, W. ., Almendros, G., Chesworth, W., Grossl, P. ., Sparks, D. L., Spaargaren, O., Fairbridge, R. ., & Singer, A. (2008). Classification of Soils: FAO. In *Encyclopedia of Soil Science* (pp. 111–113). https://link.springer.com/referenceworkentry/10.1007/978-1-4020-3995-9_102

Chow, T. E., & Hodgson, M. E. (2009). Effects of lidar post-spacing and DEM resolution to mean slope estimation. *International Journal of Geographical Information Science*, 23(10), 1277–1295. https://doi.org/10.1080/13658810802344127

Christie, D., & Neill, S. P. (2022). Measuring and Observing the Ocean Renewable Energy Resource. *Comprehensive Renewable Energy*, 1–8, 149–175. https://doi.org/10.1016/B978-0-12-819727-1.00083-2

Climent, F., Momparler, A., & Carmona, P. (2019). Anticipating bank distress in the Eurozone: An Extreme Gradient Boosting approach. *Journal of Business Research*, 101, 885–896. https://doi.org/10.1016/J.JBUSRES.2018.11.015

Cowling, R. M., MacDonald, I. A. W. W., & Simmons, M. T. (1996). The Cape Peninsula, South Africa: Physiographical, biological and historical background to an extraordinary hot-spot of biodiversity. *Biodiversity and Conservation*, 5(5), 527–550. https://doi.org/10.1007/BF00137608/METRICS

Cowling, R. M., Richardson, D. M., Shultze, R. E., Hoffmann, M. T., Midgley, J. J., & Hilton-Taylor, C. (1966). Species richness at the regional scale. In R. M. Cowling, D. M. Richardson, & S. M. Pierce (Eds.), *Vegetation of Southern Africa*. Cambridge University Press.

Cui, Z., Qing, X., Chai, H., Yang, S., Zhu, Y., & Wang, F. (2021). Real-time rainfall-runoff prediction using light gradient boosting machine coupled with singular spectrum analysis. *Journal of Hydrology*, 603, 127124. https://doi.org/10.1016/J.JHYDROL.2021.127124

Das, T. (2022). *Google Colab: Everything you Need to Know*. Geekflare.

https://geekflare.com/google-colab/

Dávila-Hernández, S., González-Trinidad, J., Júnez-Ferreira, H. E., Bautista-Capetillo, C. F., Morales de Ávila, H., Cázares Escareño, J., Ortiz-Letechipia, J., Robles Rovelo, C. O., & López-Baltazar, E. A. (2022). Effects of the Digital Elevation Model and Hydrological Processing Algorithms on the Geomorphological Parameterization. *Water*, 14(15), 2363. https://doi.org/10.3390/W14152363

De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., & Crombé, P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, 186, 39–49. https://doi.org/10.1016/J.GEOMORPH.2012.12.015

Del Valle, H., Tentor, F., Sione, W., Zamboni, P., Acenolaza, P., & Metternicht, G. (2022). Vertical Accuracy Assessment of Freely Available Digital Elevation Models: Implications for Low-Relief Landscapes. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2022-July, 6414–6417. https://doi.org/10.1109/IGARSS46834.2022.9884896

Department of Environment and Science Queensland. (2020). *Terrain roughness*. https://wetlandinfo.des.qld.gov.au/wetlands/ecology/aquatic-ecosystems-natural/estuarine-marine/itst/terrain-roughness/

DFFE. (2023). *SA National Land-Cover Datasets | Environmental Geographical Information Systems (E-GIS)*. https://egis.environment.gov.za/sa_national_land_cover_datasets

Dikpal, R. L., Renuka Prasad, T. J., & Satish, K. (2017). Evaluation of morphometric parameters derived from Cartosat-1 DEM using remote sensing and GIS techniques for Budigere Amanikere watershed, Dakshina Pinakini Basin, Karnataka, India. *Applied Water Science*, 7(8), 4399–4414. https://doi.org/10.1007/S13201-017-0585-6/TABLES/6

Ding, Y., Chen, Z., Lu, W., & Wang, X. (2021). A CatBoost approach with wavelet decomposition to improve satellite-derived high-resolution PM2.5 estimates in Beijing-Tianjin-Hebei. *Atmospheric Environment*, 249, 118212. https://doi.org/10.1016/J.ATMOSENV.2021.118212

Dorogush, A. V., Ershov, V., & Yandex, A. G. (2018). *CatBoost: gradient boosting with categorical features support*. https://arxiv.org/abs/1810.11363v1

Doshi-Velez, F., & Kim, B. (2017). *Towards A Rigorous Science of Interpretable Machine Learning*. https://doi.org/10.3390/e24050687

Dosilovic, F. K., Brcic, M., & Hlupic, N. (2018). Explainable artificial intelligence: A survey. *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings*, 210–215. https://doi.org/10.23919/MIPRO.2018.8400040

Drakenstein Municipality. (2009). *River Environmental Management Plan Berg River and major tributaries within the Drakenstein Municipality of the Western Cape*. http://www.drakenstein.gov.za/docs/Documents/River EMP Part 1 Report Structure.pdf

NGBoost: Natural Gradient Boosting for Probabilistic Prediction, PartF16814 37th International Conference on Machine Learning, ICML 2020 2670 (2020).

https://arxiv.org/abs/1910.03225v4

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1, 100002. https://doi.org/10.1016/J.SRS.2020.100002

DWS. (2002). *Water Resources Management Plan in the Diep River Catchment: A Situation Assessment* . https://www.dws.gov.za/iwqs/reports/Diep/Water_Resources_Management_Plan_in_the_Diep_River_Catchment_A_Situation_Assessment.pdf

DWS. (2020). *Department of Water and Sanitation South Africa: Part 1 of 3. Government Gazette*.

Echeverria, V., Gabriela, A., Larreta, E., Crespo, P. R., & Mulas, M. (2022). Lineament Extraction from Digital Terrain Derivate Model: A Case Study in the Girón–Santa Isabel Basin, South Ecuador. *Remote Sensing*, 14(21), 5400. https://doi.org/10.3390/RS14215400

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813. https://doi.org/10.1111/J.1365-2656.2008.01390.X

Emery, W., & Camps, A. (2017). Orbital Mechanics, Image Navigation, and Cartographic Projections. *Introduction to Satellite Remote Sensing*, 565–596. https://doi.org/10.1016/B978-0-12-809254-5.00007-5

Epuh, E. E., Okolie, C. J., Daramola, O. E., Ogunlade, F. S., Oyatayo, F. J., Akinnusi, S. A., & Emmanuel, E. O. I. (2020). An integrated lineament extraction from satellite imagery and gravity anomaly maps for groundwater exploration in the Gongola Basin. *Remote Sensing Applications: Society and Environment*, 20, 100346. https://doi.org/10.1016/J.RSASE.2020.100346

Esch, T., Bachofer, F., Heldens, W., Hirner, A., Marconcini, M., Palacios-Lopez, D., Roth, A., Üreyen, S., Zeidler, J., Dech, S., & Gorelick, N. (2018). Where We Live—A Summary of the Achievements and Planned Evolution of the Global Urban Footprint. *Remote Sensing*, 10(6), 895. https://doi.org/10.3390/RS10060895

Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., & Strano, E. (2017). Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 30–42. https://doi.org/10.1016/J.ISPRSJPRS.2017.10.012

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenbock, H., Muller, A., & Dech, S. (2013). Urban footprint processor-Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. *IEEE Geoscience and Remote Sensing Letters*, 10(6), 1617–1621. https://doi.org/10.1109/LGRS.2013.2272953

Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felbier, A., Thiel, M., Schmidt, M., Müller, A., & Dech, S. (2012). TanDEM-X mission—new perspectives for the inventory and

monitoring of global settlement patterns. *Journal of Applied Remote Sensing*, 6(1), 61701–61702. https://doi.org/10.1117/1.JRS.6.061702

Esch, T., Thiel, M., Schenk, A., Roth, A., Müller, A., & Dech, S. (2010). Delineation of Urban footprints from TerraSAR-X data by analyzing speckle characteristics and intensity information. *IEEE Transactions on Geoscience and Remote Sensing*, 48(2), 905–916. https://doi.org/10.1109/TGRS.2009.2037144

ESRI. (2021a). *How Point to Raster works*. https://desktop.arcgis.com/en/arcmap/latest/tools/conversion-toolbox/point-to-raster.htm

ESRI. (2021b). *Spatial Autocorrelation (Global Moran's I)*. ESRI - ArcGIS Desktop. https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-statistics-toolbox/spatial-autocorrelation.htm

ESRI. (2023). *How Spatial Autocorrelation (Global Moran's I) works*. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm

Ettritch, G., Hardy, A., Bojang, L., Cross, D., Bunting, P., & Brewer, P. (2018). Enhancing digital elevation models for hydraulic modelling using flood frequency detection. *Remote Sensing of Environment*, 217, 506–522. https://doi.org/10.1016/J.RSE.2018.08.029

Fan, J., Wang, X., Zhang, F., Ma, X., & Wu, L. (2020). Predicting daily diffuse horizontal solar radiation in various climatic regions of China using support vector machine and tree-based soft computing models with local and extrinsic climatic data. *Journal of Cleaner Production*, 248, 119264. https://doi.org/10.1016/J.JCLEPRO.2019.119264

Fauvel, M., Chanussot, J., & Benediktsson, J. A. (2006). Decision fusion for the classification of urban remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10), 2828–2838. https://doi.org/10.1109/TGRS.2006.876708

Feng, S., Lin, Y., Wang, Y., Yang, Y., Shen, W., Teng, F., & Hong, W. (2020). DEM Generation With a Scale Factor Using Multi-Aspect SAR Imagery Applying Radargrammetry. *Remote Sensing*, 12(3), 556. https://doi.org/10.3390/RS12030556

Ferré, J. (2009). Regression Diagnostics. *Comprehensive Chemometrics*, 3, 33–89. https://doi.org/10.1016/B978-044452701-1.00076-4

Florinsky, I. V. (2010). Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science*, 12(1), 47–62. https://doi.org/10.1080/136588198242003

Forthofer, R. N., Lee, E. S., & Hernandez, M. (2007). Linear Regression. *Biostatistics*, 349–386. https://doi.org/10.1016/B978-0-12-369492-8.50018-2

Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1), 119–139. https://doi.org/10.1006/JCSS.1997.1504

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Ann. Statist*, 29(5), 1189–1232. https://doi.org/10.1214/AOS/1013203451

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view

of boosting (With discussion and a rejoinder by the authors). *Ann. Statist.*, 28(2), 337–407. https://doi.org/10.1214/AOS/1016218223

Fu, C. Y., Tsay, J.-R. R., Roth, A., Knöpfle, W., Strunz, G., Lehner, M., Reinartz, P., Gruen, A., Akca, D., Cohen, L., Cohen, I., Kass, M., Witkin, A., Terzopoulos, D., Podobnikar, T., Bohinj, Fu, C. Y., Tsay, J.-R. R., Fuss, C., … Kressner, D. (2016). STATISTIC TESTS AIDED MULTI-SOURCE DEM FUSION. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41(July), 227–233. https://doi.org/10.5194/isprsarchives-xli-b6-227-2016

Gallant, J. C., & Wilson, J. P. (2000). Primary topographic attributes. *Terrain Analysis: Principles and Applications*, *January*, 51–85.

Gao, J. (2010). Impact of sampling intervals on the reliability of topographic variables mapped from grid DEMs at a micro-scale. *International Journal of Geographical Information Science*, 12(8), 875–890. https://doi.org/10.1080/136588198241545

Gao, W., Zhou, L., Liu, S., Guan, Y., Gao, H., & Hui, B. (2022). Machine learning prediction of lignin content in poplar with Raman spectroscopy. *Bioresource Technology*, 348, 126812. https://doi.org/10.1016/J.BIORTECH.2022.126812

Gdulová, K., Marešová, J., & Moudrý, V. (2020). Accuracy assessment of the global TanDEM-X digital elevation model in a mountain environment. *Remote Sensing of Environment*, 241, 111724. https://doi.org/10.1016/j.rse.2020.111724

Gebreyesus, Y., Dalton, D., Nixon, S., Chiara, D. De, & Chinnici, M. (2023). Machine Learning for Data Center Optimizations: Feature Selection Using Shapley Additive exPlanation (SHAP). *Future Internet*, 15(3), 88. https://doi.org/10.3390/FI15030088

Gens, R., & Genderen, J. (1996). SAR interferometry : issues, techniques, applications. *International Journal of Remote Sensing*, 17, 1803–1835. https://doi.org/10.1080/01431169608948741

Georganos, S., Grippa, T., Gadiaga, A., Vanhuysse, S., Kalogirou, S., Lennert, M., & Linard, C. (2019). An application of geographical random forests for population estimation in Dakar, Senegal using very-high-resolution satellite imagery. *2019 Joint Urban Remote Sensing Event, JURSE 2019*. https://doi.org/10.1109/JURSE.2019.8809049

Gesch, D. B., Oimoen, M. J., & Evans, G. A. (2014). Accuracy assessment of the U.S. Geological Survey National Elevation Dataset, and comparison with other large-area elevation datasets: SRTM and ASTER. *Open-File Report*, 18. https://doi.org/10.3133/OFR20141008

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42. https://doi.org/10.1007/S10994-006-6226-1/METRICS

Geymen, A. (2014). Digital elevation model (DEM) generation using the SAR interferometry technique. *Arabian Journal of Geosciences*, 7(2), 827–837. https://doi.org/10.1007/S12517-012-0811-3/FIGURES/9

Ghamisi, P., Rasti, B., Yokoya, N., Wang, Q., Hofle, B., Bruzzone, L., Bovolo, F., Chi, M., Anders, K., Gloaguen, R., Atkinson, P. M., & Benediktsson, J. A. (2019). Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of

the art. *IEEE Geoscience and Remote Sensing Magazine*, 7(1), 6–39. https://doi.org/10.1109/MGRS.2018.2890023

Ghassemian, H. (2001a). A retina based multi-resolution image fusion. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2, 709–711. https://doi.org/10.1109/IGARSS.2001.976610

Ghassemian, H. (2001b). Multi-sensor image fusion using multirate filter banks. *IEEE International Conference on Image Processing*, 1, 846–849. https://doi.org/10.1109/ICIP.2001.959178

Ghassemian, H. (2016). A review of remote sensing image fusion methods. *Information Fusion*, 32, 75–89. https://www.sciencedirect.com/science/article/pii/S1566253516300173

Gill, M., Anderson, R., Hu, H., Bennamoun, M., Petereit, J., Valliyodan, B., Nguyen, H. T., Batley, J., Bayer, P. E., & Edwards, D. (2022). Machine learning models outperform deep learning models, provide interpretation and facilitate feature selection for soybean trait prediction. *BMC Plant Biology*, 22(1), 1–8. https://doi.org/10.1186/S12870-022-03559-Z/TABLES/5

Girohi, P., & Bhardwaj, A. (2022). A Neural Network-Based Fusion Approach for Improvement of SAR Interferometry-Based Digital Elevation Models in Plain and Hilly Regions of India. *AI*, 3(4), 820–843. https://doi.org/10.3390/AI3040050

Global 2010 Tree Cover (30 m) | GLAD, (2023). https://glad.umd.edu/dataset/global-2010-tree-cover-30-m

GLAD. (2023). *Global 2010 Bare Ground (30 m)*. https://glad.umd.edu/dataset/global-2010-bare-ground-30-m

Goldblum, M., Finzi, M., Rowan, K., & Wilson, A. G. (2023). *The No Free Lunch Theorem, Kolmogorov Complexity, and the Role of Inductive Biases in Machine Learning*. https://arxiv.org/abs/2304.05366v1

Golden Software. (2023). *Surfer® | 2D & 3D mapping, modeling & analysis software for scientists and engineers*. https://www.goldensoftware.com/products/surfer

Google. (2023). *Google Colab*. https://research.google.com/colaboratory/faq.html

Government of Canada. (2023). *High Resolution Digital Elevation Model (HRDEM) - CanElevation Series - Open Government Portal*. https://open.canada.ca/data/en/dataset/957782bf-847c-4644-a757-e383c0057995

Gqomfa, B., Maphanga, T., & Shale, K. (2022). The impact of informal settlement on water quality of Diep River in Dunoon. *Sustainable Water Resources Management*, 8(1), 1–18. https://doi.org/10.1007/S40899-022-00629-W/TABLES/2

Graczyk, M., Lasota, T., Trawiński, B., & Trawiński, K. (2010). Comparison of bagging, boosting and stacking ensembles applied to real estate appraisal. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5991 LNAI(PART 2), 340–350. https://doi.org/10.1007/978-3-642-12101-2_35/COVER

GRASS-Development-Team. (2023). *r.terrain.texture - GRASS GIS manual*. https://grass.osgeo.org/grass82/manuals/addons/r.terrain.texture.html

Grohmann, C. H. (2015). Effects of spatial resolution on slope and aspect derivation for regional-scale analysis. *Computers & Geosciences*, 77, 111–117. https://doi.org/10.1016/J.CAGEO.2015.02.003

Grohmann, C. H., Smith, M. J., & Riccomini, C. (2011). Multiscale analysis of topographic surface roughness in the Midland Valley, Scotland. *IEEE Transactions on Geoscience and Remote Sensing*, 49(4), 1200–1213. https://doi.org/10.1109/TGRS.2010.2053546

Guo, J. M., Liu, Y. F., Chang, C. H., & Nguyen, H. S. (2012). Improved hand tracking system. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(5), 693–701. https://doi.org/10.1109/TCSVT.2011.2177192

Guth, P. L., & Geoffroy, T. M. (2021). LiDAR point cloud and ICESat-2 evaluation of 1 second global digital elevation models: Copernicus wins. *Transactions in GIS*, 25(5), 2245–2261. https://doi.org/10.1111/TGIS.12825

Guth, P. L., Van Niekerk, A., Grohmann, C. H., Muller, J. P., Hawker, L., Florinsky, I. V., Gesch, D., Reuter, H. I., Herrera-Cruz, V., Riazanoff, S., López-Vázquez, C., Carabajal, C. C., Albinet, C., & Strobl, P. (2021). Digital Elevation Models: Terminology and Definitions. *Remote Sensing*, 13(18), 3581. https://doi.org/10.3390/RS13183581

Habtemariam, E. T., Kekeba, K., Martínez-Ballesteros, M., & Martínez-Álvarez, F. (2023). A Bayesian Optimization-Based LSTM Model for Wind Power Forecasting in the Adama District, Ethiopia. *Energies*, 16(5), 2317. https://doi.org/10.3390/EN16052317

Hafeez, M. ., Rashid, M., Tariq, H., Abideen, Z. U., Alotaibi, S. ., & Sinky, M. . (2021). Performance Improvement of Decision Tree: A Robust Classifier Using Tabu Search Algorithm. *Applied Sciences*, 11(15). https://doi.org/10.3390/app11156728

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1), 6–23. https://doi.org/10.1109/5.554205

Han, J., Fang, M., Ye, S., Chen, C., Wan, Q., & Qian, X. (2019). Using Decision Tree to Predict Response Rates of Consumer Satisfaction, Attitude, and Loyalty Surveys. *Sustainability*, 11(8), 2306. https://doi.org/10.3390/SU11082306

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. *Journal of Big Data*, 7(1), 1–45. https://doi.org/10.1186/S40537-020-00369-8/FIGURES/9

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850–853. https://doi.org/10.1126/SCIENCE.1244693/SUPPL_FILE/HANSEN.SM.PDF

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. https://doi.org/10.1007/978-0-387-84858-7

Haugerud, R. A., & Harding, D. J. (2001). Some algorithms for virtual deforestation of lidar topographic survey data. *IAPRS 22-24 October, Annapolis*, 211–218.

Hawker, L., Neal, J., & Bates, P. (2019). Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. *Remote Sensing of Environment*, 232, 111319. https://doi.org/10.1016/J.RSE.2019.111319

Hawker, L., Rougier, J., Neal, J., Bates, P., Archer, L., & Yamazaki, D. (2018). Implications of Simulating Global Digital Elevation Models for Flood Inundation Studies. *Water Resources Research*, 54, 7910–7928. https://doi.org/10.1029/2018WR023279

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., & Neal, J. (2022). A 30 m global map of elevation with forests and buildings removed. *Environmental Research Letters*, 17(2), 24016. https://doi.org/10.1088/1748-9326/AC4D4F

Henderson, D. ., Ferrill, D. ., & Clarke, K. . (1996). Mapping geological faults using image processing techniques applied to hill-shaded digital elevation models. *Proceeding of Southwest Symposium on Image Analysis and Interpretation, San Antonio, TX, USA*, 240–245. https://doi.org/10.1109/IAI.1996.493760

Hengl, T., Heuvelink, G. B. M., & Van Loon, E. E. (2010). On the uncertainty of stream networks derived from elevation data: The error propagation approach. *Hydrology and Earth System Sciences*, 14(7), 1153–1165. https://doi.org/10.5194/HESS-14-1153-2010

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832–844. https://doi.org/10.1109/34.709601

Hobson, R. D. (2019). Surface roughness in topography: Quantitative approach. *Spatial Analysis in Geomorphology*, 6, 221–245. https://doi.org/10.4324/9780429273346-8/SURFACE-ROUGHNESS-TOPOGRAPHY-QUANTITATIVE-APPROACH-HOBSON

Höhle, J., & Höhle, M. (2009). Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(4), 398–406. https://doi.org/10.1016/J.ISPRSJPRS.2009.02.003

Hoja, D., Reinartz, P., & Schroeder, M. (2006). Comparison of DEM Generation and Combination Methods Using High Resolution Optical Stereo Imagery and Interferometric SAR Data. *Revue Francaise de Photogrammetrie et de Teledetection*, *184*, 89–94. https://www.semanticscholar.org/paper/8296f905520fab744b4aa24f41054065471a9569

Hoos, H. H. (2010). *Programming by Optimisation*.

Horton, R. . (1945). Erosional Development of Streams and Their Drainage Basins; Hydrophysical Approach to Quantitative Morphology. *Geological Society of America Bulletin*, 56, 275–370. https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/89983

Horton, R. E. (1932). Drainage-basin characteristics. *Eos, Transactions American Geophysical Union*, 13(1), 350–361. https://doi.org/10.1029/TR013I001P00350

Hssina, B., Merbouha, A., Ezzikouri, H., & Erritali, M. (n.d.). A comparative study of decision tree ID3 and C4.5. *IJACSA) International Journal of Advanced Computer Science and Applications, Special Issue on Advances in Vehicular Ad Hoc Networking and Applications*, 13. www.ijacsa.thesai.org

Hu, M., & Ji, S. (2022). Accuracy evaluation and improvement of common DEM in Hubei Region based on ICESat/GLAS data. *Earth Science Informatics*, 15(1), 221–231. https://doi.org/10.1007/S12145-021-00721-3/TABLES/2

Hu, Y. (2003). *Automated Extraction of Digital Terrain Models, Roads and Buildings Using Airborne Lidar Data* [University of Calgary]. http://hdl.handle.net/1880/40305

Huang, H., Wu, D., Fang, L., & Zheng, X. (2022). Comparison of Multiple Machine Learning Models for Estimating the Forest Growing Stock in Large-Scale Forests Using Multi-Source Data. *Forests*, 13(9), 1471. https://doi.org/10.3390/F13091471

Huang, S. (2023). Linear regression analysis. *International Encyclopedia of Education: Fourth Edition*, 548–557. https://doi.org/10.1016/B978-0-12-818630-5.10067-3

Huber, M., Osterkamp, N., Marschalk, U., Tubbesing, R., Wendleder, A., Wessel, B., & Roth, A. (2021). Shaping the Global High-Resolution TanDEM-X Digital Elevation Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 7198–7212. https://doi.org/10.1109/JSTARS.2021.3095178

Hussain, S., Mustafa, M. W., Jumani, T. A., Baloch, S. K., Alotaibi, H., Khan, I., & Khan, A. (2021). A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection. *Energy Reports*, 7, 4425–4436. https://doi.org/10.1016/J.EGYR.2021.07.008

Incorporating Boosted Regression Trees into Ecological Latent Variable Models, Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (2011). www.eBird.org

Huy Thong, P., Hoang Son, L., & Dinh Hoa, N. (2012). A PARALLEL RANDOM FOREST ALGORITHM FOR DIGITAL ELEVATION MODEL CLASSIFICATION. *International Symposium on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences*.

Ibrahim, S. (2023). Improving Land Use/Cover Classification Accuracy from Random Forest Feature Importance Selection Based on Synergistic Use of Sentinel Data and Digital Elevation Model in Agriculturally Dominated Landscape. *Agriculture*, 13(1), 98. https://doi.org/10.3390/AGRICULTURE13010098

IGN. (2018). *RGE ALTI ® Version 2.0 Descriptif de contenu*.

Iheaturu, C. J., Ayodele, E. G., & Okolie, C. J. (2020). An assessment of the accuracy of structure-from-motion (SfM) photogrammetry for 3D terrain mapping. *Geomatics, Landmanagement and Landscape*, no. 2, 65–82. https://doi.org/10.15576/GLL/2020.2.65

Iwahashi, J., & Pike, R. J. (2007). Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology*, 86(3–4), 409–440. https://doi.org/10.1016/J.GEOMORPH.2006.09.012

Jabeur, S. Ben, Gharib, C., Mefteh-Wali, S., & Arfi, W. Ben. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, 166, 120658. https://doi.org/10.1016/J.TECHFORE.2021.120658

Jain, S. ., & Singh, V. . (2003). Emerging Techniques for Data Acquisition and Systems Modeling. *Developments in Water Science*, 51(C), 123–205.

https://doi.org/10.1016/S0167-5648(03)80057-6

Javed, A. R., Ahmed, W., Pandya, S., Maddikunta, P. K. R., Alazab, M., & Gadekallu, T. R. (2023). A Survey of Explainable Artificial Intelligence for Smart Cities. *Electronics*, 12(4), 1020. https://doi.org/10.3390/ELECTRONICS12041020

Dataset | ALOS@EORC, (2023). https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm

Jimenez, L. O., & Morales-Morell, A. (1999). Classification of hyperdimensional data based on feature and decision fusion approaches using projection pursuit, majority voting, and neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3 I), 1360–1366. https://doi.org/10.1109/36.763300

Johnson, R., & Zhang, T. (2014). Learning nonlinear functions using regularized greedy forest. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(5), 942–954. https://doi.org/10.1109/TPAMI.2013.159

Jordan, G. (2007). Digital terrain analysis in a GIS environment. Concepts and development. *Lecture Notes in Geoinformation and Cartography*, 0(9783540367307), 1–43. https://doi.org/10.1007/978-3-540-36731-4_1/COVER

Joseph, M. (2020). *The Gradient Boosters II: Regularized Greedy Forest – Deep & Shallow*. https://deep-and-shallow.com/2020/02/09/the-gradient-boosters-ii-regularized-greedy-forest/

Kadavi, P. R., Lee, C. W., & Lee, S. (2018). Application of Ensemble-Based Machine Learning Models to Landslide Susceptibility Mapping. *Remote Sensing*, 10(8), 1252. https://doi.org/10.3390/RS10081252

Kadra, A., Lindauer, M., Hutter, F., & Grabocka, J. (2021). Well-tuned Simple Nets Excel on Tabular Datasets. *Advances in Neural Information Processing Systems*, 29, 23928–23941. https://arxiv.org/abs/2106.11189v2

Kashifi, M. T., & Ahmad, I. (2022). Efficient Histogram-Based Gradient Boosting Approach for Accident Severity Prediction With Multisource Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(6), 236–258. https://doi.org/10.1177/03611981221074370

Kasi, V., Yeditha, P. K., Rathinasamy, M., Pinninti, R., Landa, S. R., Sangamreddi, C., Agarwal, A., & Dandu Radha, P. R. (2020). A novel method to improve vertical accuracy of CARTOSAT DEM using machine learning models. *Earth Science Informatics*, 13(4), 1139–1150. https://doi.org/10.1007/S12145-020-00494-1/TABLES/3

Kavzoglu, T., & Teke, A. (2022). Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arabian Journal for Science and Engineering*, 47(6), 7367–7385. https://doi.org/10.1007/S13369-022-06560-8/FIGURES/12

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems*, 30. https://github.com/Microsoft/LightGBM.

Kearns, M. J., & Valiant, L. G. Cryptographic limitations on learning Boolean formulae and finite automata. *Journal of the ACM (JACM)*, 41(1). https://link.springer.com/chapter/10.1007/3-540-56483-7_21

Khramtsov, V., Sergeyev, A., Spiniello, C., Tortora, C., Napolitano, N. R., Agnello, A., Getman, F., De Jong, J. T. A., Kuijken, K., Radovich, M., Shan, H. Y., & Shulga, V. (2019). KiDS-SQuaD - II. Machine learning selection of bright extragalactic objects to search for new gravitationally lensed quasars. *Astronomy & Astrophysics*, 632, A56. https://doi.org/10.1051/0004-6361/201936006

Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS*, 8(1), 83–111. https://doi.org/10.1111/J.1467-9671.2004.00169.X

Kim, D.-E., Gourbesville, P., & Liong, S.-Y. (2019). Overcoming data scarcity in flood hazard assessment using remote sensing and artificial neural network. *Smart Water 2019 4:1*, 4(1), 1–15. https://doi.org/10.1186/S40713-018-0014-5

Kim, D. E., Liong, S. Y., Gourbesville, P., Andres, L., & Liu, J. (2020). Simple-Yet-Effective SRTM DEM Improvement Scheme for Dense Urban Cities Using ANN and Remote Sensing Data: Application to Flood Modeling. *Water*, 12(3), 816. https://doi.org/10.3390/W12030816

Kim, D. E., Liu, J., Liong, S. Y., Gourbesville, P., & Strunz, G. (2021). Satellite DEM Improvement Using Multispectral Imagery and an Artificial Neural Network. *Water*, 13(11), 1551. https://doi.org/10.3390/W13111551

Explainable AI (XAI) Methods Part 1 — Partial Dependence Plot (PDP) | by Seungjun (Josh) Kim | Towards Data Science, (2021). https://towardsdatascience.com/explainable-ai-xai-methods-part-1-partial-dependence-plot-pdp-349441901a3d

Klemelii, J., Klinke, S., & Sofyan, H. (2000). Classification and Regression Trees. *XploRe® - Application Guide*, 281–304. https://doi.org/10.1007/978-3-642-57292-0_10

Knöpfle, W., Strunz, G., & Roth, A. (1998). *Mosaicking of Digital Elevation Models derived by SAR Interferometry*. https://semanticscholar.org/paper/fbeb44e27d73246575c220ff6ecb4fa8c2f85057

Koehn, D., Lessmann, S., & Schaal, M. (2020). Predicting online shopping behaviour from clickstream data using deep learning. *Expert Systems with Applications*, 150, 113342. https://doi.org/10.1016/J.ESWA.2020.113342

Kolesnikov, A. A., Kikin, P. M., & Portnov, A. M. (2019). DISEASES SPREAD PREDICTION IN TROPICAL AREAS BY MACHINE LEARNING METHODS ENSEMBLING AND SPATIAL ANALYSIS TECHNIQUES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3-W8(3/W8), 221–226. https://doi.org/10.5194/ISPRS-ARCHIVES-XLII-3-W8-221-2019

Kölsch, M., & Turk, M. (2004). Robust hand detection. *Proceedings - Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 614–619. https://doi.org/10.1109/AFGR.2004.1301601

Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. *Informatica*, 31(3), 249–268. https://www.informatica.si/index.php/informatica/article/view/148

Kotsiantis, S. B. (2013). Decision trees: A recent overview. *Artificial Intelligence Review*, 39(4), 261–283. https://doi.org/10.1007/S10462-011-9272-4/METRICS

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M., & Zink, M. (2007). TanDEM-X: A satellite formation for high-resolution SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11), 3317–3340. https://doi.org/10.1109/TGRS.2007.900693

Krieger, G., Zink, M., Bachmann, M., Bräutigam, B., Schulze, D., Martone, M., Rizzoli, P., Steinbrecher, U., Antony, J., Zan, F., Hajnsek, I., Papathanassiou, K., Kugler, F., Cassola, M., Younis, M., Baumgartner, S., López-Dekker, P., Prats, P., Moreira, A., … Moreira, A. (2013). TanDEM-X: A radar interferometer with two formation-flying satellites. *Acta Astronautica*, 89, 83–98. https://doi.org/10.1016/J.ACTAASTRO.2013.03.008

Kulp, S., & Strauss, B. (2018). CoastalDEM: A global coastal digital elevation model improved from SRTM using a neural network. *Remote Sensing of Environment*, 206, 231–239. https://doi.org/10.1016/J.RSE.2017.12.026

Kuncheva, L. I. (2004). Combining Pattern Classifiers. *Combining Pattern Classifiers*. https://doi.org/10.1002/0471660264

Lane, S. N. (2005). Roughness – time for a re-evaluation? *Earth Surface Processes and Landforms*, 30(2), 251–253. https://doi.org/10.1002/ESP.1208

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003

Lary, D. J., Zewdie, G. K., Liu, X., Wu, D., Levetin, E., Allee, R. J., Malakar, N., Walker, A., Mussa, H., Mannino, A., & Aurin, D. (2018). Machine Learning Applications for Earth Observation. *Earth Observation Open Science and Innovation*, 165–218. https://doi.org/10.1007/978-3-319-65633-5_8

Laurini, R., & Thompson, D. (1992). Geometries: Position, representation, dimensions. *Fundamentals of Spatial Information Systems*, 113–174. https://doi.org/10.1016/B978-0-08-092420-5.50009-8

Leenaers, H., & Okx, J. P. (1989). The use of digital elevation models for flood hazard mapping. *Earth Surface Processes and Landforms*, 14(6), 631–640. https://doi.org/10.1002/ESP.3290140617

Li, H., Zhao, J., Yan, B., Yue, L., & Wang, L. (2022). Global DEMs vary from one to another: an evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data. *International Journal of Digital Earth*, 15(1), 1149–1168. https://doi.org/10.1080/17538947.2022.2094002

Li, K., Xu, H., & Liu, X. (2022). Analysis and visualization of accidents severity based on LightGBM-TPE. *Chaos, Solitons & Fractals*, 157, 111987. https://doi.org/10.1016/J.CHAOS.2022.111987

Li, S., Kang, X., Fang, L., Hu, J., & Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. *Information Fusion*, 33, 100–112. https://doi.org/10.1016/j.inffus.2016.05.004

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy*, 23(1), 18. https://doi.org/10.3390/E23010018

Liu, W., Deng, K., Zhang, X., Cheng, Y., Zheng, Z., Jiang, F., & Peng, J. (2020). A Semi-Supervised Tri-CatBoost Method for Driving Style Recognition. *Symmetry*, 12(3), 336. https://doi.org/10.3390/SYM12030336

Liu, Y., Aleksandrov, M., Zlatanova, S., Zhang, J., Mo, F., & Chen, X. (2019). Classification of Power Facility Point Clouds from Unmanned Aerial Vehicles Based on Adaboost and Topological Constraints. *Sensors*, 19(21), 4717. https://doi.org/10.3390/S19214717

Liu, Y., Bates, P. D., Neal, J. C., & Yamazaki, D. (2021). Bare-Earth DEM Generation in Urban Areas for Flood Inundation Simulation Using Global Digital Elevation Models. *Water Resources Research*, 57(4), e2020WR028516. https://doi.org/10.1029/2020WR028516

Liu, Y., Chen, X., Yang, J., Li, L., & Wang, T. (2023). Snow avalanche susceptibility mapping from tree-based machine learning approaches in ungauged or poorly-gauged regions. *CATENA*, 224, 106997. https://doi.org/10.1016/J.CATENA.2023.106997

Lodha, S. K., Fitzpatrick, D. M., & Helmbold, D. P. (2007). Aerial lidar data classification using AdaBoost. *3DIM 2007 - Proceedings 6th International Conference on 3-D Digital Imaging and Modeling*, 435–442. https://doi.org/10.1109/3DIM.2007.10

Loh, W. Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 14–23. https://doi.org/10.1002/WIDM.8

Louppe, G., & Geurts, P. (2012). Ensembles on random patches. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7523 LNAI(PART 1), 346–361. https://doi.org/10.1007/978-3-642-33460-3_28/COVER

Lu, Y., Fan, X., Zhang, Y., Wang, Y., & Jiang, X. (2023). Machine Learning Models Using SHapley Additive exPlanation for Fire Risk Assessment Mode and Effects Analysis of Stadiums. *Sensors*, 23(4), 2151. https://doi.org/10.3390/S23042151

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 2017-Decem, 4766–4775. https://arxiv.org/abs/1705.07874v2

Ma, Y., Liu, H., Jiang, B., Meng, L., Guan, H., Xu, M., Cui, Y., Kong, F., Yin, Y., & Wang, M. (2020). An Innovative Approach for Improving the Accuracy of Digital Elevation Models for Cultivated Land. *Remote Sensing*, 12(20), 3401.

https://doi.org/10.3390/RS12203401

Malindi, M. (2022). *On the development of a digital elevation model over South Africa using ground and satellite data* [University of Cape Town]. https://open.uct.ac.za/bitstream/handle/11427/37482/thesis_ebe_2022_malindi mihlali.pdf?sequence=4&isAllowed=y

Mangolini, M. (1994). *Apport de la fusion d'images satellitaires multicapteurs au niveau pixel en télédétection et photo-interprétation*. Université de Nice Sophia-Antipolis.

Mariani, S., & Sipper, M. (2022). High Per Parameter: A Large-Scale Study of Hyperparameter Tuning for Machine Learning Algorithms. *Algorithms 2022, Vol. 15, Page 315*, 15(9), 315. https://doi.org/10.3390/A15090315

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., … Zwally, J. (2017). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment*, 190, 260–273. https://doi.org/10.1016/J.RSE.2016.12.029

Martinez-Cantin, R., De Freitas, N., Doucet, A., & Castellanos, J. A. (2008). Active Policy Learning for Robot Planning and Exploration under Uncertainty. *Robotics: Science and Systems*, 3, 321–328. https://doi.org/10.7551/MITPRESS/7830.003.0041

Maune, D. F. (2001). *Digital Elevation Model Technologies and Applica_tions: The DEM Users Manual*.

Maune, D. F. (2007). *Digital Elevation Model Technologies and Applications: The DEM User Manual*. https://doi.org/1-57083-082-7 (2007)

McCarty, D. A., Kim, H. W., & Lee, H. K. (2020). Evaluation of Light Gradient Boosted Machine Learning Technique in Large Scale Land Use and Land Cover Classification. *Environments*, 7(10), 84. https://doi.org/10.3390/ENVIRONMENTS7100084

Meadows, M. E., & Compton, J. S. (2015). Table Mountain: Wonder of Nature at the Foot of Africa. *World Geomorphological Landscapes*, 95–102. https://doi.org/10.1007/978-3-319-03560-4_11/COVER

Meadows, M., & Wilson, M. (2021). A Comparison of Machine Learning Approaches to Improve Free Topography Data for Flood Modelling. *Remote Sensing*, 13(2), 275. https://doi.org/10.3390/RS13020275

Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for data fusion. *Information Fusion*, 57(2), 115–129. https://doi.org/10.1016/j.inffus.2019.12.001

Merry, C. L. (2009). EGM2008 evaluation for Africa. *Newon's Bulletin*, 200–206. https://www.isgeoid.polimi.it/Newton/Newton_4/Report_EA10_Africa.pdf (Date accessed:

Mesa-Mingorance, J. L., & Ariza-López, F. J. (2020). Accuracy Assessment of Digital Elevation Models (DEMs): A Critical Review of Practices of the Past Three Decades. *Remote Sensing*, 12(16), 2630. https://doi.org/10.3390/RS12162630

Miao, X., Heaton, J. S., Zheng, S., Charlet, D. A., & Liu, H. (2011). Applying tree-based ensemble algorithms to the classification of ecological zones using multi-temporal multi-source remote-sensing data. *International Journal of Remote Sensing*, 33(6), 1823–1849. https://doi.org/10.1080/01431161.2011.602651

Michalopoulou, M., Depountis, N., Nikolakopoulos, K., & Boumpoulis, V. (2022). The Significance of Digital Elevation Models in the Calculation of LS Factor and Soil Erosion. *Land*, 11(9), 1592. https://doi.org/10.3390/LAND11091592

Microsoft Corporation. (2023). *Parameters*. LightGBM. https://lightgbm.readthedocs.io/en/latest/Parameters.html

Minár, J., & Sládek, J. (2009). Morphological network as an indicator of a morphotectonic field in the central Western Carpathians (Slovakia). *Zeitschrift Für Geomorphologie, Supplementary Issues*, 53(2), 23–29. https://doi.org/10.1127/0372-8854/2009/0053S3-0023

Mitrentsis, G., & Lens, H. (2022). An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting. *Applied Energy*, 309, 118473. https://doi.org/10.1016/J.APENERGY.2021.118473

Mohammadi, A., Karimzadeh, S., Jalal, S. J., Kamran, K. V., Shahabi, H., Homayouni, S., & Al-Ansari, N. (2020). A Multi-Sensor Comparative Analysis on the Suitability of Generated DEM from Sentinel-1 SAR Interferometry Using Statistical and Hydrological Models. *Sensors*, 20(24), 7214. https://doi.org/10.3390/S20247214

Molnar, C., König, G., Herbinger, J., Freiesleben, T., Dandl, S., Scholbeck, C. A., Casalicchio, G., Grosse-Wentrup, M., & Bischl, B. (2022). General Pitfalls of Model-Agnostic Interpretation Methods for Machine Learning Models. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13200 LNAI, 39–68. https://doi.org/10.1007/978-3-031-04083-2_4/FIGURES/7

Moore, I., Grayson, R., & Ladson, A. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5, 3–30. https://doi.org/10.1002/HYP.3360050103

Moser, K., Ahn, C., & Noe, G. (2007). Characterization of microtopography and its influence on vegetation patterns in created wetlands. *Wetlands*, 27(4), 1081–1097. https://doi.org/10.1672/0277-5212(2007)27[1081:COMAII]2.0.CO;2

Mphuthi, M. S. (2021). *Towards a geoid consistent vertical datum in South Africa*. University of Cape Town.

Mück, M., Klotz, M., & Taubenbock, H. (2017). Validation of the DLR Global Urban Footprint in rural areas: A case study for Burkina Faso. *2017 Joint Urban Remote Sensing Event, JURSE 2017*. https://doi.org/10.1109/JURSE.2017.7924618

Mura, M. D., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., & Benediktsson, J. A. (2015). *Challenges and Opportunities of Multimodality and Data Fusion in Remote Sensing*. 103(9).

Muthusamy, M., Casado, M. R., Butler, D., & Leinster, P. (2021). Understanding the effects

of Digital Elevation Model resolution in urban fluvial flood modelling. *Journal of Hydrology*, 596, 126088. https://doi.org/10.1016/J.JHYDROL.2021.126088

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7(DEC). https://doi.org/10.3389/FNBOT.2013.00021/BIBTEX

Nauta, M., Trienes, J., Nguyen, E., Peters, M., Schmitt, Y., Schlötterer, J., Van Keulen, M., Seifert, C., Trienes, J., Schmitt, Y., Schlötterer, J., Seifert, C., Nguyen, E., Peters, M., van Keulen, M., Pathak, S., & van Keulen, M. (2023). From Anecdotal Evidence to Quantitative Evaluation Methods: A Systematic Review on Evaluating Explainable AI. *ACM Computing Surveys*, 55(13s). https://doi.org/10.1145/3583558

ngboost · PyPI. (n.d.). https://pypi.org/project/ngboost/

Nguyen, N.-H., & Woo, D.-M. (2014). Application of Multi-Class AdaBoost Algorithm to Terrain Classification of Satellite Images. *Journal of IKEEE*, 18(4), 536–543. https://doi.org/10.7471/IKEEE.2014.18.4.536

Nie, X., Guo, W., Huang, B., Zhuo, M., Li, D., Li, Z., & Yuan, Z. (2019). Effects of soil properties, topography and landform on the understory biomass of a pine forest in a subtropical hilly region. *CATENA*, 176, 104–111. https://doi.org/10.1016/J.CATENA.2019.01.007

Noh, M.-J. J., & Howat, I. M. (2015). Automated stereo-photogrammetric DEM generation at high latitudes: Surface Extraction with TIN-based Search-space Minimization (SETSM) validation and demonstration over glaciated regions. *GIScience \& Remote Sensing*, 52(2), 198–217. https://doi.org/10.1080/15481603.2015.1008621

Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2022). Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Computer Methods and Programs in Biomedicine*, 214, 106584. https://doi.org/10.1016/J.CMPB.2021.106584

Nohara, Y., Wakata, Y., & Nakashima, N. (2015). Interpreting Medical Information Using Machine Learning and Individual Conditional Expectation. *Studies in Health Technology and Informatics*, 216, 1073. https://doi.org/10.3233/978-1-61499-564-7-1073

Nordin, N., Zainol, Z., Mohd Noor, M. H., & Chan, L. F. (2023). An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. *Asian Journal of Psychiatry*, 79, 103316. https://doi.org/10.1016/J.AJP.2022.103316

Nwilo, P. C., Ayodele, E. G., & Okolie, C. J. (2017). Determination of the Impacts of Landscape Offsets on the 30-metre SRTM DEM Through a Comparative Analysis with Bare-Earth Elevations. *FIG Peer Review Journal*, 21p. https://www.fig.net/resources/proceedings/fig_proceedings/fig2017/papers/ts06h/TS06H_nwilo_ayodele_et_al_8560.pdf

Nwilo, P. C., Ogbeta, C. O., Daramola, O. E., Okolie, C. J., & Orji, M. J. (2021). Soil Erosion Susceptibility Mapping of Imo River Basin Using Modified Geomorphometric Prioritisation Method. *Quaestiones Geographicae*, 40(3), 143–162. https://doi.org/10.2478/QUAGEO-2021-0029

Nwilo, P. C., Olayinka, D. N., Okolie, C. J., Emmanuel, E. I., Orji, M. J., & Daramola, O. E. (2020). Impacts of land cover changes on desertification in northern Nigeria and implications on the Lake Chad Basin. *Journal of Arid Environments*, 181, 104190. https://doi.org/10.1016/J.JARIDENV.2020.104190

Nwilo, P. C., Onyegbula, J. C., Okolie, C. J., Daramola, O. E., Abolaji, O. E., & Arungwa, I. D. (2022). Influence of land cover, slope, and aspect on the vertical accuracy of SPOT DEM at selected sites in Nigeria. *Applied Geomatics*, 14(1), 17–31. https://doi.org/10.1007/S12518-021-00404-0/TABLES/6

Nyuytiymbiy, K. (2020). *Parameters and Hyperparameters in Machine Learning and Deep Learning*. Towards Data Science. https://towardsdatascience.com/parameters-and-hyperparameters-aa609601a9ac

O'Leary, D. ., & Friedman, J. . (1978). Towards a workable lineament symbology. *Proceedings of the Third International Conference on the New Basement Tectonics, Basement Tectonics Committee Publication #3*, 29–31.

O'leary, D. W., Friedman, J. D., & Pohn, H. A. (1976). Lineament, linear, lineation: Some proposed new standards for old terms: Discussion. *Bull. Geol. Soc. Am.*, 89, 1463–1469.

O'Loughlin, F., Paiva, R. C. D., Durand, M., Alsdorf, D., & Bates, P. (2016). A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sensing of Environment*, 182, 49–59. https://doi.org/10.1016/J.RSE.2016.04.018

O'Neil, G. L., Saby, L., Band, L. E., & Goodall, J. L. (2019). Effects of LiDAR DEM Smoothing and Conditioning Techniques on a Topography-Based Wetland Identification Model. *Water Resources Research*, 55(5), 4343–4363. https://doi.org/10.1029/2019WR024784

Okolie, C. J., & Arungwa, I. D. (2022). Accuracy assessment of ALOS World 3D-30m DEMs using field data from Lagos and Federal Capital Territory, Nigeria. *Https://Doi.Org/10.1080/10106049.2022.2060320*, 37(26), 11686–11705. https://doi.org/10.1080/10106049.2022.2060320

Okolie, C. J., & Smit, J. L. (2022). A systematic review and meta-analysis of Digital elevation model (DEM) fusion: pre-processing, methods and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188, 1–29. https://doi.org/10.1016/J.ISPRSJPRS.2022.03.016

Olajubu, V., Trigg, M. A., Berretta, C., Sleigh, A., Chini, M., Matgen, P., Mojere, S., & Mulligan, J. (2021). Urban correction of global DEMs using building density for Nairobi, Kenya. *Earth Science Informatics*, 14(3), 1383–1398. https://doi.org/10.1007/S12145-021-00647-W/TABLES/5

Olusina, J. O., & Okolie, C. J. (2018). Visualisation of Uncertainty in 30m Resolution Global Digital Elevation Models: SRTM v3.0 and ASTER v2. *Nigerian Journal of Technological Development*, 15(3), 77. https://doi.org/10.4314/njtd.v15i3.2

Omer, Z. M., & Shareef, H. (2022). Comparison of decision tree based ensemble methods for prediction of photovoltaic maximum current. *Energy Conversion and Management: X*, 16, 100333. https://doi.org/10.1016/J.ECMX.2022.100333

Ong, B., Lee, T. M., Li, G., & Chuen, D. L. K. (2015). Evaluating the Potential of Alternative Cryptocurrencies. *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data*, 81–135. https://doi.org/10.1016/B978-0-12-802117-0.00005-9

Ordnance Survey. (2023). *OS Terrain 5 | Data Products | Ordnance Survey*. https://www.ordnancesurvey.co.uk/products/os-terrain-5

Orimoloye, I. R., Ololade, O. O., Mazinyo, S. P., Kalumba, A. M., Ekundayo, O. Y., Busayo, E. T., Akinsanola, A. A., & Nel, W. (2019). Spatial assessment of drought severity in Cape Town area, South Africa. *Heliyon*, 5(7), e02148. https://doi.org/10.1016/J.HELIYON.2019.E02148

Ortega, M. (2020). *Burning down the black box of ML using SHAP*. Vantage AI - Data Science. https://www.vantage-ai.com/en/blog/burning-down-the-black-box-of-ml-using-shap

Owino, R. . (2021). *EVALUATING THE POTENTIAL OF OPENSOURCE GEOSPATIAL DATA IN TOPOGRAPHICAL MAP REVISION*. University of Nairobi.

Ozturk, M. Y., & Colkesen, I. (2021). EVALUATION OF EFFECTIVENESS OF PATCH BASED IMAGE CLASSIFICATION TECHNIQUE USING HIGH RESOLUTION WORLDVIEW-2 IMAGE. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-4-W5-2021(4/W5-2021), 417–423. https://doi.org/10.5194/ISPRS-ARCHIVES-XLVI-4-W5-2021-417-2021

Padhi, D. K., Padhy, N., Kumar Bhoi, A., Shafi, J., Fazal Ijaz, M., Padhi, D. K. ;, Padhy, N. ;, Bhoi, A. K. ;, Shafi, J. ;, Ijaz, M. F. A., Popescu, G., Mosavi, A., & Adrangi, B. (2021). A Fusion Framework for Forecasting Financial Market Direction Using Enhanced Ensemble Models and Technical Indicators. *Mathematics*, 9(21), 2646. https://doi.org/10.3390/MATH9212646

Pakoksung, K., & Takagi, M. (2015). Digital elevation models on accuracy validation and bias correction in vertical. *Modeling Earth Systems and Environment*, 2, 1–13. https://doi.org/10.1007/s40808-015-0069-3

Palar, P. S., Zuhal, L. R., & Shimoyama, K. (2023). Enhancing the explainability of regression-based polynomial chaos expansion by Shapley additive explanations. *Reliability Engineering & System Safety*, 232, 109045. https://doi.org/10.1016/J.RESS.2022.109045

Pang, Y., Lefsky, M., Sun, G., Miller, M., & Li, Z. (2008). TEMPERATE FOREST HEIGHT ESTIMATION PERFORMANCE USING ICESAT GLAS DATA FROM DIFFERENT OBSERVATION PERIODS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing 2008*. https://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=5A9E1937F42BEC9A551F49DEDAA6D35C?doi=10.1.1.183.9950&rep=rep1&type=pdf

Papadopoulos, S., Azar, E., Woon, W. L., & Kontokosta, C. E. (2017). Evaluation of tree-based ensemble learning algorithms for building energy performance estimation. *Journal of Building Performance Simulation*, 11(3), 322–332. https://doi.org/10.1080/19401493.2017.1354919

Papasaika-Hanusch, C. (2012). *Doctor of Sciences presented by* (Issue 20845).

Papasodoro, C., Royer, A., Langlois, A., & Berthier, E. (2016). Potential of RADARSAT-2 stereo radargrammetry for the generation of glacier DEMs. *Journal of Glaciology*, 62, 486–496. https://doi.org/10.1017/JOG.2016.44

Pavlis, N. K., Holmes, S. A., Kenyon, S. C., & Factor, J. K. (2012). The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). *Journal of Geophysical Research: Solid Earth*, 117(B4). https://doi.org/10.1029/2011JB008916

PCI - Geomatics. (2022). *CATALYST Professional*. https://catalyst.earth/products/catalyst-pro/

Pepe, A., & Calò, F. (2017). A Review of Interferometric Synthetic Aperture RADAR (InSAR) Multi-Track Approaches for the Retrieval of Earth's Surface Displacements. *Applied Sciences*, 7(12), 1264. https://doi.org/10.3390/APP7121264

Pepple, G. T., Bala, A., Aro-Lambo, Y. A., & Ibukun, O. J. (2023). Map Revision of Small Scaled Topographic Sheet 303 Abakaliki South_West (SW), Nigeria. *FIG Working Week 2023 Protecting Our World, Conquering New Frontiers Orlando, Florida, USA, 28 May–1 June 2023*, 17p. https://www.fig.net/resources/proceedings/fig_proceedings/fig2023/papers/cinema01/CINEMA01_pepple_bala_et_al_11979.pdf

Petrasova, A., Mitasova, H., Petras, V., & Jeziorska, J. (2017). Fusion of high-resolution DEMs for water flow modeling. *Open Geospatial Data, Software and Standards*, 2(1), 0–7. https://doi.org/10.1186/s40965-017-0019-2

Peucker, T. K., Fowler, R. J., Little, J. J., & Mark, D. M. (1978). The triangulated irregular network. *Proceedings of the ASP Digital Terrain Models (DTM) Symposium. Falls Church, Virginia.*, 516–540.

Pham, H. T., Marshall, L., Johnson, F., & Sharma, A. (2018). A method for combining SRTM DEM and ASTER GDEM2 to improve topography estimation in regions without reference data. *Remote Sensing of Environment*, 210(June 2017), 229–241. https://doi.org/10.1016/J.RSE.2018.03.026

Pham, X. T. T., & Ho, T. H. (2021). Using boosting algorithms to predict bank failure: An untold story. *International Review of Economics & Finance*, 76, 40–54. https://doi.org/10.1016/J.IREF.2021.05.005

Pike, R. J., Evans, I. S., & Hengl, T. (2009). Chapter 1 Geomorphometry: A Brief Guide. *Developments in Soil Science*, 33(C), 3–30. https://doi.org/10.1016/S0166-2481(08)00001-9

Pipaud, I., Loibl, D., & Lehmkuhl, F. (2015). Evaluation of TanDEM-X elevation data for geomorphological mapping and interpretation in high mountain environments — A case study from SE Tibet, China. *Geomorphology*, 246, 232–254. https://doi.org/10.1016/J.GEOMORPH.2015.06.025

Pittman, S. J., & Brown, K. A. (2011). Multi-Scale Approach for Predicting Fish Species Distributions across Coral Reef Seascapes. *PLOS ONE*, 6(5), e20583. https://doi.org/10.1371/JOURNAL.PONE.0020583

Podhorányi, M., Unucka, J., Bobál', P., & Říhová, V. (2013). Effects of LIDAR DEM resolution in hydrodynamic modelling: model sensitivity for cross-sections. *International*

Journal of Digital Earth, 6(1), 3–27. https://doi.org/10.1080/17538947.2011.596578

Podobnikar, T. (2005). Production of integrated digital terrain model from multiple datasets of different quality. *International Journal of Geographical Information Science*, 19(1), 69–89. https://doi.org/10.1080/13658810412331280130

Pohl, C., & van Genderen, J. (2016). Remote Sensing Image Fusion : A Practical Guide. *Remote Sensing Image Fusion*. https://doi.org/10.1201/9781315370101

Pohl, C., & Van Genderen, J. L. (1998). Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5), 823–854. https://doi.org/10.1080/014311698215748

Polidori, L., & El Hage, M. (2020). Methods : A Critical Review. *Remote Sensing*, 12(21), 3522.

Polidori, L., & Hage, M. El. (2020). Digital Elevation Model Quality Assessment Methods: A Critical Review. *Remote Sensing*, 12(21), 3522. https://doi.org/10.3390/RS12213522

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 6(3), 21–44. https://doi.org/10.1109/MCAS.2006.1688199

Prasad, R. K., Mondal, N. C., Banerjee, P., Nandakumar, M. V., & Singh, V. S. (2008). Deciphering potential groundwater zone in hard rock through the application of GIS. *Environmental Geology*, 55(3), 467–475. https://doi.org/10.1007/S00254-007-0992-3/FIGURES/9

Preety, K., Prasad, A. K., Varma, A. K., & El-Askary, H. (2022). Accuracy Assessment, Comparative Performance, and Enhancement of Public Domain Digital Elevation Models (ASTER 30 m, SRTM 30 m, CARTOSAT 30 m, SRTM 90 m, MERIT 90 m, and TanDEM-X 90 m) Using DGPS. *Remote Sensing*, 14(6), 1334. https://doi.org/10.3390/RS14061334/S1

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 2018-Decem, 6638–6648. https://github.com/catboost/catboost

Punmiya, R., & Choe, S. (2019). Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. *IEEE Transactions on Smart Grid*, 10(2), 2326–2329. https://doi.org/10.1109/TSG.2019.2892595

Purinton, B., & Bookhagen, B. (2017). Validation of digital elevation models (DEMs) and comparison of geomorphic metrics on the southern Central Andean Plateau. *Earth Surface Dynamics*, 5(2), 211–237. https://doi.org/10.5194/ESURF-5-211-2017

Purinton, B., & Bookhagen, B. (2021). Beyond Vertical Point Accuracy: Assessing Inter-pixel Consistency in 30 m Global DEMs for the Arid Central Andes. *Frontiers in Earth Science*, 9, 901. https://doi.org/10.3389/FEART.2021.758606/BIBTEX

Python. (2023). *time — Time access and conversions*¶. The Python Standard Library » Generic Operating System Services. https://docs.python.org/3/library/time.html

QGIS. (2023). *Welcome to the QGIS project!* https://qgis.org/en/site/

Qi-yong, Y., Zhong-cheng, J., Wen-jun, L., & Hui, L. (2014). Prediction of soil organic matter in peak-cluster depression region using kriging and terrain indices. *Soil and Tillage Research*, 144, 126–132. https://doi.org/10.1016/J.STILL.2014.07.011

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106. https://doi.org/10.1007/BF00116251

Raggam, J., Gutjahr, K., Perko, R., & Schardt, M. (2010). Assessment of the Stereo-Radargrammetric Mapping Potential of TerraSAR-X Multibeam Spotlight Data. *IEEE Transactions on Geoscience and Remote Sensing*, 48, 971–977. https://doi.org/10.1109/TGRS.2009.2037315

Rao, B. S., Kumar, G. A., Gopala krishna, P. V. S. S. N., Srinivasulu, P., & Venkataraman, V. R. (2012). Evaluation of EGM 2008 with EGM96 and its Utilization in Topographical Mapping Projects. *Journal of the Indian Society of Remote Sensing*, 40(2), 335–340. https://doi.org/10.1007/S12524-011-0131-1/TABLES/4

RDocumentation. (2023). *lm: Fitting Linear Models*. https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/lm

Reuter, H. I., Hengl, T., Gessler, P., & Soille, P. (2009). Chapter 4 Preparation of DEMs for Geomorphometric Analysis. *Developments in Soil Science*, 33(C), 87–120. https://doi.org/10.1016/S0166-2481(08)00004-4

Reuter, H. I., Nelson, A., Strobl, P., Mehl, W., & Jarvis, A. (2009). A first assessment of aster GDEM tiles for absolute accuracy, relative accuracy and terrain parameters. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 5, 240–243. https://doi.org/10.1109/IGARSS.2009.5417688

Rewicki, F., Denzler, J., & Niebling, J. (2023). *IS IT WORTH IT? COMPARING SIX DEEP AND CLASSICAL METHODS FOR UNSUPERVISED ANOMALY DETECTION IN TIME SERIES*.

Riley, S. ., Degloria, S. ., & Elliot, S. . (1999). A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity. *Intermountain Journal of Sciences*, 5(1–4), 23–27. https://download.osgeo.org/qgis/doc/reference-docs/Terrain_Ruggedness_Index.pdf

Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Borla Tridon, D., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M., Wessel, B., Krieger, G., Zink, M., & Moreira, A. (2017). Generation and performance assessment of the global TanDEM-X digital elevation model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132, 119–139. https://doi.org/10.1016/J.ISPRSJPRS.2017.08.008

Rocha, J., Duarte, A., Silva, M., Fabres, S., Vasques, J., Revilla-Romero, B., & Quintela, A. (2020). The Importance of High Resolution Digital Elevation Models for Improved Hydrological Simulations of a Mediterranean Forested Catchment. *Remote Sensing*, 12(20), 3287. https://doi.org/10.3390/RS12203287

Rockinger, O. (1996). *Pixel-Level Fusion of Image Sequences using Wavelet Frames 1*.

RockWare. (2023). *Rockware*. https://www.rockware.com/

Roe, B. P., Yang, H. J., Zhu, J., Liu, Y., Stancu, I., & McGregor, G. (2005). Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nuclear*

Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 543(2–3), 577–584. https://doi.org/10.1016/J.NIMA.2004.12.018

Rokach, L., & Maimon, O. (2005). Top-down induction of decision trees classifiers - A survey. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 35(4), 476–487. https://doi.org/10.1109/TSMCC.2004.843247

Roman, D. R., Ming WANG, Y., & Saleh, J. (2010). *TS 1C-Geodetic Infrastructure and Datum Geodesy, Geoids, and Vertical Datums: A Perspective from the U.S. National Geodetic Survey*. 11–16.

Ross Quinlan, by J., Kaufmann Publishers, M., & Salzberg, S. L. (1994). *Programs for Machine Learning*. 16, 235–240.

Roth, A. E. (1988). Introduction to the Shapley value. *The Shapley Value*, 1–28. https://doi.org/10.1017/CBO9780511528446.002

Ruiyao, K., Tao, X. I. E., Ming, M. A., & Ruilin, K. (2022). Application of CatBoost model in water depth inversion. *Bulletin of Surveying and Mapping*, 0(7), 33. https://doi.org/10.13474/J.CNKI.11-2246.2022.0199

Saeed, W., & Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Systems*, 263, 110273. https://doi.org/10.1016/J.KNOSYS.2023.110273

Safaei, N., Safaei, B., Seyedekrami, S., Talafidaryani, M., Masoud, A., Wang, S., Li, Q., & Moqri, M. (2022). E-CatBoost: An efficient machine learning framework for predicting ICU mortality using the eICU Collaborative Research Database. *PLOS ONE*, 17(5), e0262895. https://doi.org/10.1371/JOURNAL.PONE.0262895

Salah, M. (2021). SRTM DEM correction over dense urban areas using inverse probability weighted interpolation and Sentinel-2 multispectral imagery. *Arabian Journal of Geosciences*, 14(9), 1–16. https://doi.org/10.1007/S12517-021-07148-6/FIGURES/15

Saleem, R., Yuan, B., Kurugollu, F., Anjum, A., & Liu, L. (2022). Explaining deep neural networks: A survey on the global interpretation methods. *Neurocomputing*, 513, 165–180. https://doi.org/10.1016/J.NEUCOM.2022.09.129

Salinas-Melgoza, M. A., Skutsch, M., & Lovett, J. C. (2018). Predicting aboveground forest biomass with topographic variables in human-impacted tropical dry forest landscapes. *Ecosphere*, 9(1), e02063. https://doi.org/10.1002/ECS2.2063

Samat, A., Li, E., Du, P., Liu, S., & Xia, J. (2021). GPU-Accelerated CatBoost-Forest for Hyperspectral Image Classification Via Parallelized mRMR Ensemble Subspace Feature Selection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 3200–3214. https://doi.org/10.1109/JSTARS.2021.3063507

Samsonov, T. (2022). Granularity of Digital Elevation Model and Optimal Level of Detail in Small-Scale Cartographic Relief Presentation. *Remote Sensing*, 14(5), 1270. https://doi.org/10.3390/RS14051270

SAPPINGTON, J. M., LONGSHORE, K. M., & THOMPSON, D. B. (2007). Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep

in the Mojave Desert. *J. of Wildlife Management*, 71(5), 1419–1426. https://doi.org/10.2193/2005-723

Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197–227. https://doi.org/10.1007/BF00116037

Schapire, R. E. (2003). The Boosting Approach to Machine Learning An Overview. *Nonlinear Estimation and Classification*. www.research.att.com/

Schindler, K., PAPASAIKA-HANUSCH, H., Schütz, S., & Baltsavias, E. (2011). Improving Wide-Area DEMs Through Data Fusion – Chances and Limits. *Photogrammetric Week '11*, 159–170. https://semanticscholar.org/paper/18dd9d1877241c4d25821b035d667324e56af6e7

Schlaffer, S., Matgen, P., Hollaus, M., & Wagner, W. (2015). Flood detection from multi-temporal SAR data using harmonic analysis and change detection. *International Journal of Applied Earth Observation and Geoinformation*, 38, 15–24. https://doi.org/10.1016/J.JAG.2014.12.001

Schmitt, M., & Zhu, X. X. (2016). Data Fusion and Remote Sensing: An ever-growing relationship. *IEEE Geoscience and Remote Sensing Magazine*, 4(4), 6–23. https://doi.org/10.1109/MGRS.2016.2561021

Schneider, P., & Xhafa, F. (2022). Anomaly detection: Concepts and methods. *Anomaly Detection and Complex Event Processing over IoT Data Streams*, 49–66. https://doi.org/10.1016/B978-0-12-823818-9.00013-4

Schumm, S. A. (1956). Evolution of Drainage Systems and Slopes in Badlands at Perth Amboy, New Jersey. *GSA Bulletin*, 67, 464–597. https://doi.org/10.1130/0016-7606(1956)67[597:EODSAS]2.0.CO;2

Schwalbe, G., Finzel, B., Schwalbe, G., & Finzel, B. (2023). A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Mining and Knowledge Discovery 2023*, 1–59. https://doi.org/10.1007/S10618-022-00867-8

Scikit-Learn. (2023a). *scikit-learn: machine learning in Python — scikit-learn 1.3.0 documentation*. https://scikit-learn.org/stable/

Scikit-Learn. (2023b). *sklearn.ensemble.BaggingRegressor*. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.BaggingRegressor.html#sklearn.ensemble.BaggingRegressor

Scikit-Learn. (2023c). *sklearn.ensemble.GradientBoostingRegressor*. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & De Freitas, N. (2016). Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1), 148–175. https://doi.org/10.1109/JPROC.2015.2494218

Shen, K., Qin, H., Zhou, J., & Liu, G. (2022). Runoff Probability Prediction Model Based on Natural Gradient Boosting with Tree-Structured Parzen Estimator Optimization. *Water*, 14(4), 545. https://doi.org/10.3390/W14040545

Shen, S. S. (1990). Summary of types of data fusion methods utilized in workshop papers. *Multisource Data Integration in Remote Sensing, Proceedings of Workshop, Maryland, USA*, 145–149.

Shetty, S., Umesh, P., & Shetty, A. (2021). Lineament Extraction from Open-Source Digital Elevation Models: a Comparative Analysis. *2021 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, 66–71. https://doi.org/10.1109/DISCOVER52564.2021.9663691

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81, 84–90. https://doi.org/10.1016/J.INFFUS.2021.11.011

Sidek, O., & Quadri, S. A. (2012). A review of data fusion models and systems. *International Journal of Image and Data Fusion*, *1*, 3–21. https://doi.org/10.1080/19479832.2011.645888

Simmons, M. T., & Cowling, R. M. (1996). Why is the Cape Peninsula so rich in plant species? An analysis of the independent diversity components. *Biodiversity and Conservation*, 5(5), 551–573. https://doi.org/10.1007/BF00137609/METRICS

Smith, L. C. (2010). Emerging Applications of Interferometric Synthetic Aperture Radar (InSAR) in Geomorphology and Hydrology. *Annals of the Association of American Geographers*, 92(3), 385–398. https://doi.org/10.1111/1467-8306.00295

Smith, M. W. (2014). Roughness in the Earth Sciences. *Earth-Science Reviews*, 136, 202–225. https://doi.org/10.1016/J.EARSCIREV.2014.05.016

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems 25 (NIPS 2012)*.

Soliman, A., & Han, L. (2019). Effects of vertical accuracy of digital elevation model (DEM) data on automatic lineaments extraction from shaded DEM. *Advances in Space Research*, 64(3), 603–622. https://doi.org/10.1016/J.ASR.2019.05.009

Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. V. (2017). A review of machine learning techniques using decision tree and support vector machine. *Proceedings - 2nd International Conference on Computing, Communication, Control and Automation, ICCUBEA 2016*. https://doi.org/10.1109/ICCUBEA.2016.7860040

Song, H., Huang, B., Liu, Q., & Zhang, K. (2015). Improving the spatial resolution of landsat TM/ETM+ through fusion with SPOT5 images via learning-based super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3), 1195–1204. https://doi.org/10.1109/TGRS.2014.2335818

Speybroeck, N. (2012). Classification and regression trees. *International Journal of Public Health*, 57(1), 243–246. https://doi.org/10.1007/S00038-011-0315-Z/FIGURES/2

Stanford ML Group. (2023). *No Title*. NGBoost. https://github.com/stanfordmlgroup/ngboost/blob/master/ngboost/distns/normal.py

Sterkenburg, T. F., & Grünwald, P. D. (2021). The no-free-lunch theorems of supervised learning. *Synthese*, 199(3–4), 9979–10015. https://doi.org/10.1007/S11229-021-03233-1/FIGURES/2

Štular, B., Lozić, E., & Eichert, S. (2021). Airborne LiDAR-Derived Digital Elevation Model for Archaeology. *Remote Sensing*, 13(9), 1855. https://doi.org/10.3390/RS13091855

Su, Y., & Guo, Q. (2014). *A practical method for SRTM DEM correction over vegetated mountain areas*. 87, 216–228. https://doi.org/10.1016/J.ISPRSJPRS.2013.11.009

Su, Y., Guo, Q., Ma, Q., & Li, W. (2015). SRTM DEM Correction in Vegetated Mountain Areas through the Integration of Spaceborne LiDAR, Airborne LiDAR, and Optical Imagery. *Remote Sensing*, 7(9), 11202–11225. https://doi.org/10.3390/RS70911202

Sukristiyanti, S., Maria, R., & Lestiana, H. (2018). Watershed-based Morphometric Analysis: A Review. *IOP Conf. Series: Earth and Environmental Science*. https://doi.org/10.1088/1755-1315/118/1/012028

Sun, H., He, J., Chen, Y., & Zhao, B. (2021). Space-Time Sea Surface pCO2 Estimation in the North Atlantic Based on CatBoost. *Remote Sensing*, 13(14), 2805. https://doi.org/10.3390/RS13142805

Szepannek, G., & Lübke, K. (2022). Explaining Artificial Intelligence with Care: Analyzing the Explainability of Black Box Multiclass Machine Learning Models in Forensics. *KI - Kunstliche Intelligenz*, 36(2), 125–134. https://doi.org/10.1007/S13218-022-00764-8/FIGURES/6

Tabib Mahmoudi, F., Samadzadegan, F., & Reinartz, P. (2015). Object recognition based on the context aware decision-level fusion in multiviews imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1), 12–22. https://doi.org/10.1109/JSTARS.2014.2362103

Takacs, M. (2018). *Quality Assessment of Danish government geographical data using the gradient boosting algorithm Catboost*.

Takaku, J., & Tadono, T. (2017). Quality updates of "AW3D" global DSM generated from ALOS PRISM. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017-July, 5666–5669. https://doi.org/10.1109/IGARSS.2017.8128293

Takaku, J., Tadono, T., Tsutsui, K., & Ichikawa, M. (2015). Quality status of high resolution global DSM generated from ALOS PRISM. *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 3854–3857. https://doi.org/10.1109/IGARSS.2015.7326665

Talebi, H., Peeters, L. J. M. M., Otto, A., & Tolosana-Delgado, R. (2022). A Truly Spatial Random Forests Algorithm for Geoscience Data Analysis and Modelling. *Mathematical Geosciences*, 54(1), 1–22. https://doi.org/10.1007/S11004-021-09946-W/FIGURES/9

Tarquini, S., Vinci, S., Favalli, M., Doumaz, F., Fornaciai, A., & Nannipieri, L. (2012). Release of a 10-m-resolution DEM for the Italian territory: Comparison with global-coverage DEMs and anaglyph-mode exploration via the web. *Computers and Geosciences*, 38(1), 168–170. https://doi.org/10.1016/J.CAGEO.2011.04.018

Taskesen, E. (2022). *hgboost's documentation!* HGBoost. https://erdogant.github.io/hgboost/pages/html/index.html

Tavares da Costa, R., Mazzoli, P., & Bagli, S. (2019). Limitations posed by free DEMs in watershed studies: The case of river Tanaro in Italy. *Frontiers in Earth Science*, 7, 141.

https://doi.org/10.3389/FEART.2019.00141/BIBTEX

Tesema, T. A. (2021). Impact of identical digital elevation model resolution and sources on morphometric parameters of Tena watershed, Ethiopia. *Heliyon*, 7(11), e08345. https://doi.org/10.1016/J.HELIYON.2021.E08345

Tiren, S. (2010). *Lineament interpretation. Short review and methodology*. http://inis.iaea.org/Search/search.aspx?orig_q=RN:42040090

Toutin, T. (2010). Impact of Radarsat-2 SAR Ultrafine-Mode Parameters on Stereo-Radargrammetric DEMs. *IEEE Transactions on Geoscience and Remote Sensing*, 48, 3816–3823. https://doi.org/10.1109/TGRS.2010.2048715

Toutin, T. (2012). Radarsat-2 DSM Generation With New Hybrid, Deterministic, and Empirical Geometric Modeling Without GCP. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 2049–2055. https://doi.org/10.1109/TGRS.2011.2170693

Trevisan, V. (2022). *Using SHAP Values to Explain How Your Machine Learning Model Works*. Towards Data Science. https://towardsdatascience.com/using-shap-values-to-explain-how-your-machine-learning-model-works-732b3f40e137

Tsyganskaya, V., Martinis, S., Twele, A., Cao, W., Schmitt, A., Marzahn, P., & Ludwig, R. (2016). A FUZZY LOGIC-BASED APPROACH FOR THE DETECTION OF FLOODED VEGETATION BY MEANS OF SYNTHETIC APERTURE RADAR DATA. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B7, 371–378. https://doi.org/10.5194/ISPRS-ARCHIVES-XLI-B7-371-2016

UN-GGIM. (2019). *The Global Fundamental Geospatial Data Themes*.

USGS. (1997). *Standards for Digital Elevation Models, Part 2: Specifications*.

USNA. (2021). *EGM96 and EGM2008 and Other Geoid Grids*. https://www.usna.edu/Users/oceano/pguth/md_help/html/egm96.htm#:~:text=The RMS difference between EGM96,to get the geoid height

USNA. (2022). *Topographic Ruggedness Index*. https://www.usna.edu/Users/oceano/pguth/md_help/html/topo_rugged_index.htm

Üstün, A., Abbak, R. A., & Zeray Öztürk, E. (2016). Height biases of SRTM DEM related to EGM96: from a global perspective to regional practice. *Survey Review*, 50(358), 26–35. https://doi.org/10.1080/00396265.2016.1218159

Uuemaa, E., Ahi, S., Montibeller, B., Muru, M., & Kmoch, A. (2020). Vertical Accuracy of Freely Available Global Digital Elevation Models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM). *Remote Sensing*, 12(21), 3482. https://doi.org/10.3390/RS12213482

Van Nieuwenhuizen, N., Lindsay, J. B., & De Vries, B. (2021). Automated Mapping of Transportation Embankments in Fine-Resolution LiDAR DEMs. *Remote Sensing*, 13(7), 1308. https://doi.org/10.3390/RS13071308

Vernimmen, R., & Hooijer, A. (2023). New LiDAR-Based Elevation Model Shows Greatest Increase in Global Coastal Exposure to Flooding to Be Caused by Early-Stage Sea-Level

Rise. *Earth's Future*, 11(1), e2022EF002880. https://doi.org/10.1029/2022EF002880

Victoria, A. H., & Maragatham, G. (2021). Automatic tuning of hyperparameters using Bayesian optimization. *Evolving Systems*, 12(1), 217–223. https://doi.org/10.1007/S12530-020-09345-2/FIGURES/7

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1. https://doi.org/10.1109/CVPR.2001.990517

Wang, Q., Shi, W., Atkinson, P. M., & Zhao, Y. (2015). Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*, 166, 191–204. https://doi.org/10.1016/J.RSE.2015.06.003

Wang, S., Ren, Z., Wu, C., Lei, Q., Gong, W., Ou, Q., Zhang, H., Ren, G., & Li, C. (2019). DEM generation from Worldview-2 stereo imagery and vertical accuracy assessment for its application in active tectonics. *Geomorphology*, 336, 107–118. https://doi.org/10.1016/J.GEOMORPH.2019.03.016

Wang, Y., Huang, X., Ren, X., Chai, Z., & Chen, X. (2022). In-process belt-image-based material removal rate monitoring for abrasive belt grinding using CatBoost algorithm. *International Journal of Advanced Manufacturing Technology*, 123(7–8), 2575–2591. https://doi.org/10.1007/S00170-022-10341-W/FIGURES/18

Wang, Y. W., Qin, C. Z., Cheng, W. M., Zhu, A. X., Wang, Y. J., & Zhu, L. J. (2021). Automatic Crater Detection by Training Random Forest Classifiers with Legacy Crater Map and Spatial Structural Information Derived from Digital Terrain Analysis. *Annals of the American Association of Geographers*, 112(5), 1328–1349. https://doi.org/10.1080/24694452.2021.1960473

Wang, Z., Shakibi, B., Jin, L., & De Freitas, N. (2014). Bayesian Multi-Scale Optimistic Optimization. *Journal of Machine Learning Research*, 33, 1005–1014. https://arxiv.org/abs/1402.7005v1

Wang, Z., Zoghi, M., Hutter, F., Matheson, D., & De Freitas, N. (n.d.). *Bayesian Optimization in High Dimensions via Random Embeddings*.

Wangchuk, S., & Bolch, T. (2020). Mapping of glacial lakes using Sentinel-1 and Sentinel-2 data and a random forest classifier: Strengths and challenges. *Science of Remote Sensing*, 2, 100008. https://doi.org/10.1016/J.SRS.2020.100008

Washburn, M. (2017). *Digital Terrain Model Generation Using Structure from Motion: Influence of Canopy Closure and Interpolation Method on Accuracy* [Texas State University]. https://digital.library.txstate.edu/handle/10877/6603

Wazid, M., Das, A. K., Chanola, V., & Park, Y. (2022). Uniting cyber security and machine learning: Advantages, challenges and future research. *ICT Express*, 8(3), 313–321. https://doi.org/10.1016/j.icte.2022.04.007

Wechsler, S. P. (2007). Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrology and Earth System Sciences*, 11(4), 1481–1500. https://doi.org/10.5194/HESS-11-1481-2007

Weiss, A. (2001). Topographic position and landforms analysis. *ESRI User Conference*.

http://jennessent.com/downloads/TPI-poster-TNC_18x22.pdf

Welty, J. L., & Jeffries, M. I. (2018). *Western United States Ruggedness Raw Values*.

Wendi, D., Liong, S. Y., Sun, Y., & Doan, C. D. (2016). An innovative approach to improve SRTM DEM using multispectral imagery and artificial neural network. *Journal of Advances in Modeling Earth Systems*, 8(2), 691–702. https://doi.org/10.1002/2015MS000536

Wendleder, A., Felbier, A., Wessel, B., Huber, M., & Roth, A. (2016). A method to estimate long-wave height errors of SRTM C-band DEM. *IEEE Geoscience and Remote Sensing Letters*, 13(5), 696–700. https://doi.org/10.1109/LGRS.2016.2538822

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., & Roth, A. (2018a). Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. *Isprs Journal of Photogrammetry and Remote Sensing*, 139, 171–182. https://doi.org/10.1016/J.ISPRSJPRS.2018.02.017

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., & Roth, A. (2018b). Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 139, 171–182. https://doi.org/10.1016/J.ISPRSJPRS.2018.02.017

Williams, G. (2011). Decision Trees. In *Data Mining with Rattle and R*. 10.1007/978-1-4419-9890-3_11. https://doi.org/10.1007/978-1-4419-9890-3_11

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79–82. https://doi.org/10.3354/CR030079

Wilson, J., & Gallant, J. (2000). *Terrain analysis : principles and applications*. https://semanticscholar.org/paper/b42e384e7ad6d2c93e87ab37809a7d178aac7d00

Wise, S. (2000). Assessing the quality for hydrological applications of digital elevation models derived from contours. *Hydrological Processes*, 14, 1909–1929. https://doi.org/10.1002/1099-1085(20000815/30)14:11/12<1909::AID-HYP45>3.0.CO;2-6

Wise, S. M. (2007). Effect of differing DEM creation methods on the results from a hydrological model. *Computers & Geosciences*, 33(10), 1351–1365. https://doi.org/10.1016/J.CAGEO.2007.05.003

Wolf, K. L., Ahn, C., & Noe, G. B. (2011). Microtopography enhances nitrogen cycling and removal in created mitigation wetlands. *Ecological Engineering*, 37(9), 1398–1406. https://doi.org/10.1016/J.ECOLENG.2011.03.013

Wolpert, D. H. (1995). The Relationship Between PAC, the Statistical Physics Framework, the Bayesian Framework, and the VC Framework. In *The Mathematics of Generalization: The Proceedings of the SFI/CNLS Workshop on Formal Approaches to Supervised Learning*. CRC Press. https://doi.org/10.1201/9780429492525-5/RELATIONSHIP-PAC-STATISTICAL-PHYSICS-FRAMEWORK-BAYESIAN-FRAMEWORK-VC-FRAMEWORK-DAVID-WOLPERT

Wolpert, D. H. (2002). The Supervised Learning No-Free-Lunch Theorems. *Soft Computing*

and Industry, 25–42. https://doi.org/10.1007/978-1-4471-0123-9_3

World Bank Group. (2015). *DIGITAL ELEVATION MODELS. A Guidance Note on how Digital Elevation Models are created and used – includes key definitions, sample Terms of Reference and how best to plan a DEM-mission*. www.worldbank.org

Wu, S., & Nagahashi, H. (2015). Analysis of generalization ability for different AdaBoost variants based on classification and regression trees. *Journal of Electrical and Computer Engineering*, 2015. https://doi.org/10.1155/2015/835357

Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., & Hong, H. (2020). Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *CATENA*, 187, 104396. https://doi.org/10.1016/J.CATENA.2019.104396

XGBoost Developers. (2022). *No Title*. XGBoost Documentation. https://xgboost.readthedocs.io/en/stable/parameter.html

Xia, Y., He, L., Li, Y., Liu, N., & Ding, Y. (2020). Predicting loan default in peer-to-peer lending using narrative data. *Journal of Forecasting*, 39(2), 260–280. https://doi.org/10.1002/FOR.2625

Xiang, Y., Wang, F., Wan, L., & You, H. (2017). SAR-PC: Edge Detection in SAR Images via an Advanced Phase Congruency Model. *Remote Sensing*, 9(3), 209. https://doi.org/10.3390/RS9030209

Xu, J., Wu, Q., Zhang, J., & Tang, Z. (2012). Fast and accurate human detection using a cascade of boosted MS-LBP features. *IEEE Signal Processing Letters*, 19(10), 676–679. https://doi.org/10.1109/LSP.2012.2210870

Xu, K., Fang, J., Fang, Y., Sun, Q., Wu, C., & Liu, M. (2021). The Importance of Digital Elevation Model Selection in Flood Simulation and a Proposed Method to Reduce DEM Errors: A Case Study in Shanghai. *International Journal of Disaster Risk Science*, 12(6), 890–902. https://doi.org/10.1007/S13753-021-00377-Z/FIGURES/7

Yanalak, M., Musaoglu, N., Ipbuker, C., Sertel, E., & Kaya, S. (2012). DEM accuracy of high resolution satellite images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7335 LNCS(PART 3), 471–480. https://doi.org/10.1007/978-3-642-31137-6_36/COVER

Yang, H., & Bath, P. A. (2020). The Use of Data Mining Methods for the Prediction of Dementia: Evidence from the English Longitudinal Study of Aging. *IEEE Journal of Biomedical and Health Informatics*, 24(2), 345–353. https://doi.org/10.1109/JBHI.2019.2921418

Yang, Y., Han, C., Kang, X., & Han, D. (2007). An overview on pixel-level image fusion in remote sensing. *Proceedings of the IEEE International Conference on Automation and Logistics, ICAL 2007*, 2339–2344. https://doi.org/10.1109/ICAL.2007.4338968

Yang, Z., Zhang, Q., Ding, X., & Chen, W. (2020). Analysis of the Quality of Daily DEM Generation with Geosynchronous InSAR. *Engineering*, 6(8), 913–918. https://www.semanticscholar.org/paper/03926e7d038f06965ca09f34c8cf3f9cfd31ee64

Yi, H. C., You, Z. H., & Guo, Z. H. (2019). Construction and Analysis of Molecular Association Network by Combining Behavior Representation and Node Attributes.

Frontiers in Genetics, 10, 1106. https://doi.org/10.3389/FGENE.2019.01106/BIBTEX

YILMAZ, A., & ERDOĞAN, M. (2018). DESIGNING HIGH RESOLUTION COUNTRYWIDE DEM FOR TURKEY. *International Journal of Engineering and Geosciences*, 3(3), 98–108. https://doi.org/10.26833/IJEG.384822

Yu, J. H., Ge, L., & Li, X. (2014). Radargrammetry for digital elevation model generation using envisat reprocessed image and simulation image. *IEEE Geoscience and Remote Sensing Letters*, 11(9), 1589–1593. https://doi.org/10.1109/LGRS.2014.2301731

Yuan, F., Zhang, J. X., Zhang, L., & Gao, J. X. (2009). *DEM GENERATION FROM AIRBORNE LIDAR DATA*.

Zhang, J. (2010). Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, 1(1), 5–24. https://doi.org/10.1080/19479830903561035

Zhang, Y. ;, Liu, J. ;, Shen, W., Zhang, Y., Liu, J., & Shen, W. (2022). A Review of Ensemble Learning Algorithms Used in Remote Sensing Applications. *Applied Sciences*, 12(17), 8654. https://doi.org/10.3390/APP12178654

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies*, 58, 308–324. https://doi.org/10.1016/J.TRC.2015.02.019

Zhao, S. S., Cheng, W., Zhou, C., Chen, X., Zhang, S., Zhou, Z., Liu, H., & Chai, H. H. (2011). *Accuracy assessment of the ASTER GDEM and SRTM3 DEM: an example in the Loess Plateau and North China Plain of China*. 32(23), 8081–8093. https://doi.org/10.1080/01431161.2010.532176

Zhao, S., Wang, L., Cheng, W., Liu, H., & He, W. (2015). Rectification methods comparison for the ASTER GDEM V2 data using the ICESat/GLA14 data in the Lvliang mountains, China. *Environmental Earth Sciences*, 74(8), 6571–6590. https://doi.org/10.1007/S12665-015-4614-1/FIGURES/9

Zhou, C., Zhang, G., Yang, Z., Ao, M., Liu, Z., & Zhu, J. (2020). An Adaptive Terrain-Dependent Method for SRTM DEM Correction over Mountainous Areas. *IEEE Access*, 8, 130878–130887. https://doi.org/10.1109/ACCESS.2020.3009851

Zhou, L., Dang, X., Mu, H., Wang, B., & Wang, S. (2021). Cities are going uphill: Slope gradient analysis of urban expansion and its driving factors in China. *Science of The Total Environment*, 775, 145836. https://doi.org/10.1016/J.SCITOTENV.2021.145836

Zhou, Q., & Liu, X. (2008). Assessing uncertainties in derived slope and aspect from a grid dem. *Lecture Notes in Geoinformation and Cartography*, 0(199049), 279–306. https://doi.org/10.1007/978-3-540-77800-4_15/COVER

Zhu, H., Basir, O., & Karray, F. (2002). Data fusion for pattern classification via the Dempster-Shafer evidence theory. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 7, 109–110. https://doi.org/10.1109/ICSMC.2002.1175676

Zhu, Y., Liu, X., Zhao, J., Cao, J., Wang, X., & Li, D. (2019). Effect of DEM Interpolation Neighbourhood on Terrain Factors. *ISPRS International Journal of Geo-Information*, 8(1), 30. https://doi.org/10.3390/IJGI8010030

Ziadat, F. M. (2007). Effect of Contour Intervals and Grid Cell Size on the Accuracy of DEMs and Slope Derivatives. *Transactions in GIS*, 11(1), 67–81. https://doi.org/10.1111/J.1467-9671.2007.01033.X

APPENDIX 1

Bayesian-optimised hyperparameter values for XGBoost – Copernicus DEM

Urban/industrial	Agricultural	Mountain
{'target': -0.8634712331553432, 'params': {'colsample_bytree': 0.6122635019435381, 'gamma': 7.141397032504509, 'learning_rate': 0.07979845325586296, 'max_depth': 25.83753412059056, 'min_child_weight': 5.5229195062894885, 'n_estimators': 3258.562160566575, 'reg_alpha': 9.768574973376722, 'reg_lambda': 4.341970209979277, 'subsample': 0.9423466987933455}}	{'target': -0.6000511727091429, 'params': {'colsample_bytree': 0.7729197958430247, 'gamma': 0.8502501163237717, 'learning_rate': 0.028265443755881692, 'max_depth': 48.96829991527602, 'min_child_weight': 3.368925630426256, 'n_estimators': 19744.124281600318, 'reg_alpha': 2.9524074319491476, 'reg_lambda': 3.6133169469152406, 'subsample': 0.7621421392127061}}	{'target': -1.8779177165704577, 'params': {'colsample_bytree': 0.5142830149697702, 'gamma': 5.924186447163539, 'learning_rate': 0.04654576767872573, 'max_depth': 30.76969774317048, 'min_child_weight': 1.7053241844605467, 'n_estimators': 3253.5145976709905, 'reg_alpha': 9.488860483979607, 'reg_lambda': 9.656323767542286, 'subsample': 0.8084165083816495}}
Peninsula	**Grassland/shrubland**	
{'target': -1.1714332098898779, 'params': {'colsample_bytree': 0.7729197958430247, 'gamma': 0.8502501163237717, 'learning_rate': 0.028265443755881692, 'max_depth': 48.96829991527602, 'min_child_weight': 3.368925630426256, 'n_estimators': 19744.124281600318, 'reg_alpha': 2.9524074319491476, 'reg_lambda': 3.6133169469152406, 'subsample': 0.7621421392127061}}	{'target': -0.8520674927642965, 'params': {'colsample_bytree': 0.7729197958430247, 'gamma': 0.8502501163237717, 'learning_rate': 0.028265443755881692, 'max_depth': 48.96829991527602, 'min_child_weight': 3.368925630426256, 'n_estimators': 19744.124281600318, 'reg_alpha': 2.9524074319491476, 'reg_lambda': 3.6133169469152406, 'subsample': 0.7621421392127061}}	

APPENDIX 2

Bayesian-optimised hyperparameter values for LightGBM – Copernicus DEM

Urban/industrial	Agricultural	Mountain
{'target': 1.4722170960598429, 'params': {'bagging_fraction': 0.0003325043997328992, 'feature_fraction': 0.16767799568642147, 'lambda_l1': 2.3570176132046528, 'lambda_l2': 4.964866849736224, 'learning_rate': 0.015655061185629288, 'max_depth': 28.150154914297893, 'min_child_weight': 0.7578588252471351, 'min_gain_to_split': 0.8140317018005334, 'num_iterations': 12972.575127038182}}	{'target': 0.9065150342035901, 'params': {'bagging_fraction': 1.0, 'feature_fraction': 1.0, 'lambda_l1': 10.0, 'lambda_l2': 10.0, 'learning_rate': 0.0001, 'max_depth': 1.0, 'min_child_weight': 0.0001, 'min_gain_to_split': 0.0001, 'num_iterations': 9939.41449723742}}	{'target': 2.3393568003859264, 'params': {'bagging_fraction': 0.35553392626289326, 'feature_fraction': 0.24358811805970174, 'lambda_l1': 9.597333613499165, 'lambda_l2': 2.560796459484974, 'learning_rate': 0.05456540511228084, 'max_depth': 19.227792419833058, 'min_child_weight': 0.15836125810542936, 'min_gain_to_split': 0.3934480389943779, 'num_iterations': 17135.598496426985}
Peninsula	**Grassland/shrubland**	
{'target': 1.4421678184261253, 'params': {'bagging_fraction': 0.35553392626289326, 'feature_fraction': 0.24358811805970174, 'lambda_l1': 9.597333613499165, 'lambda_l2': 2.560796459484974, 'learning_rate': 0.05456540511228084, 'max_depth': 19.227792419833058, 'min_child_weight': 0.15836125810542936, 'min_gain_to_split': 0.3934480389943779, 'num_iterations': 17135.598496426985}}	{'target': 1.0916184603323762, 'params': {'bagging_fraction': 0.6489452157699611, 'feature_fraction': 0.05904137187757679, 'lambda_l1': 3.9462622914956893, 'lambda_l2': 7.984195456793634, 'learning_rate': 0.059194838241882076, 'max_depth': 47.24921501102361, 'min_child_weight': 0.9282120854294301, 'min_gain_to_split': 0.3135866328867675, 'num_iterations': 44738.48682165348}}	

APPENDIX 3

Bayesian-optimised hyperparameter values for CatBoost – Copernicus DEM

Urban/industrial	Agricultural	Mountain
{'target': 1.9120882505342984, 'params': {'bagging_temperature': 3.7916892685334296, 'border_count': 186.84179895010064, 'colsample_bylevel': 0.7498513606466211, 'depth': 14.652013197398212, 'iterations': 28644.628258570836, 'l2_leaf_reg': 4.269177292281843, 'leaf_estimation_iterations': 1.990317946206209, 'learning_rate': 0.011404571678042498, 'random_strength': 3.0058662364951974}}	{'target': 0.9208549564329822, 'params': {'bagging_temperature': 3.7916892685334296, 'border_count': 186.84179895010064, 'colsample_bylevel': 0.7498513606466211, 'depth': 14.652013197398212, 'iterations': 28644.628258570836, 'l2_leaf_reg': 4.269177292281843, 'leaf_estimation_iterations': 1.990317946206209, 'learning_rate': 0.011404571678042498, 'random_strength': 3.0058662364951974}}	{'target': 2.5333879716658902, 'params': {'bagging_temperature': 3.7916892685334296, 'border_count': 186.84179895010064, 'colsample_bylevel': 0.7498513606466211, 'depth': 14.652013197398212, 'iterations': 28644.628258570836, 'l2_leaf_reg': 4.269177292281843, 'leaf_estimation_iterations': 1.990317946206209, 'learning_rate': 0.011404571678042498, 'random_strength': 3.0058662364951974}}
Peninsula	**Grassland/shrubland**	
{'target': 1.4936981375763736, 'params': {'bagging_temperature': 3.7916892685334296, 'border_count': 186.84179895010064, 'colsample_bylevel': 0.7498513606466211, 'depth': 14.652013197398212, 'iterations': 28644.628258570836, 'l2_leaf_reg': 4.269177292281843, 'leaf_estimation_iterations': 1.990317946206209, 'learning_rate': 0.011404571678042498, 'random_strength': 3.0058662364951974}}	{'target': 1.3305591041229334, 'params': {'bagging_temperature': 3.7916892685334296, 'border_count': 186.84179895010064, 'colsample_bylevel': 0.7498513606466211, 'depth': 14.652013197398212, 'iterations': 28644.628258570836, 'l2_leaf_reg': 4.269177292281843, 'leaf_estimation_iterations': 1.990317946206209, 'learning_rate': 0.011404571678042498, 'random_strength': 3.0058662364951974}}	

APPENDIX 4

Python Algorithm for Bayesian optimisation of XGBoost

```
# Make DMatrix for calling function, to optimise speed of training

dtrain = xgb.DMatrix(X_train, label=y_train,
feature_names=X_train.columns.values)
dtest = xgb.DMatrix(X_test, label=y_test,
feature_names=X_test.columns.values)


# Define objective function

def hyp_xgb(max_depth, subsample, colsample_bytree, min_child_weight,
gamma, learning_rate, n_estimators, reg_alpha, reg_lambda):
  params = {
  'booster': 'gbtree',
  'sampling_method': 'uniform',
  'objective': 'reg:squarederror',
  'eval_metric':'rmse',
  'silent': 1
   }
  params['max_depth'] = int(round(max_depth))
  params['subsample'] = max(min(subsample, 1), 0)
  params['colsample_bytree'] = max(min(colsample_bytree, 1), 0)
  params['min_child_weight'] = (min_child_weight)
  params['gamma'] = max(gamma, 0)
  params['learning_rate'] = (learning_rate)
  params['n_estimators'] = int(n_estimators)
  params['reg_alpha'] = (reg_alpha)
  params['reg_lambda'] = (reg_lambda)
  cv_results = xgb.cv(params, dtrain, num_boost_round=1000, nfold=5,
early_stopping_rounds=10, verbose_eval=True)
  return -1.0 * cv_results['test-rmse-mean'].iloc[-1]
```

```
# Set hyperparamter bounds

pds ={
 'max_depth':(1, 50),
 'subsample':(0.0001, 1.0),
 'colsample_bytree':(0.0001, 1.0),
 'min_child_weight':(0.0001, 10),
 'gamma': (0.0001, 10),
 'learning_rate': (0.0001, 1.0),
 'n_estimators': (1, 50000),
 'reg_alpha': (0.0001, 10),
```

```
 'reg_lambda': (0.0001, 10)
 }
```

```
# Surrogate model

optimizer = BayesianOptimization(hyp_xgb, pds, verbose=2,
random_state=42)
```

```
# Run optimiser

start_time = time.time()
optimizer.maximize(init_points=10, n_iter=50)
print("--- %s seconds ---" % (time.time() - start_time))
```

```
# The best combination of parameters and target value found can be
accessed via the property optimizer.max.

optimizer.max
```

```
# While the list of all parameters probed and their corresponding
target values is available via the property optimizer.res.
for i, res in enumerate(optimizer.res):
  print("Iteration {}: \n\t{}".format(i, res))
```

APPENDIX 5

Python Algorithm for Bayesian optimisation of LightGBM

```
dtrain = lgb.Dataset(data=X_train, label=y_train)
dtest = lgb.Dataset(data=X_test, label=y_test)
```

```
# Define objective function

def hyp_lgb(num_iterations, learning_rate, feature_fraction,
bagging_fraction, max_depth, min_gain_to_split, min_child_weight,
lambda_l1, lambda_l2):

    params = {'application': 'regression',
         'early_stopping_round': 10,
         'boosting_type': 'gbdt',
         'metric':'rmse',
         } # Default parameters
    params['num_iterations'] = int(num_iterations)
    params['learning_rate'] = learning_rate
    params['feature_fraction'] = max(min(feature_fraction, 1), 0)
    params['bagging_fraction'] = max(min(bagging_fraction, 1), 0)
    params['max_depth'] = int(round(max_depth))
    params['min_gain_to_split'] = min_gain_to_split
    params['min_child_weight'] = min_child_weight
    params['lambda_l1'] = lambda_l1
    params['lambda_l2'] = lambda_l2
    cv_results = lgb.cv(params, dtrain, nfold=5, seed=42,
categorical_feature=[], stratified=False, verbose_eval = True,
early_stopping_rounds=10, num_boost_round=1000)
    return np.max(cv_results['rmse-mean'])
```

```
# Set hyperparamter bounds

pds = {'num_iterations': (1, 50000), 'learning_rate': (0.0001, 1),
     'feature_fraction': (0.0001, 1),
     'bagging_fraction': (0.0001, 1),
     'max_depth': (1, 50),
     'min_gain_to_split': (0.0001, 1),
     'min_child_weight': (0.0001, 1),
     'lambda_l1': (0.0001, 10),
     'lambda_l2': (0.0001, 10),
      }
```

```
# Surrogate model

optimizer = BayesianOptimization(hyp_lgb, pds, verbose=2,
random_state=42)
```

```
# Run optimiser

start_time = time.time()
optimizer.maximize(init_points=10, n_iter=50)
print("--- %s seconds ---" % (time.time() - start_time))
```

```
# The best combination of parameters and target value found can be
accessed via the property optimizer.max.

optimizer.max
```

```
# While the list of all parameters probed and their corresponding
target values is available via the property optimizer.res.
for i, res in enumerate(optimizer.res):
  print("Iteration {}: \n\t{}".format(i, res))
```

APPENDIX 6

Python Algorithm for Bayesian optimisation of CatBoost

```
cat_features = []

cv_dataset = cgb.Pool(data=X_train,
          label=y_train,
          cat_features=cat_features)
```

```
# Define objective function

def hyp_cat(depth, bagging_temperature, learning_rate, iterations,
l2_leaf_reg, random_strength, border_count, colsample_bylevel,
leaf_estimation_iterations):
  params = {
        "boosting_type": "Plain",
        "eval_metric": "RMSE",
        "loss_function": "RMSE",
        "verbose": True,
        "use_best_model": True,
        "random_state": 42}
  params["depth"] = int(round(depth))
  params["bagging_temperature"] = (bagging_temperature)
  params["learning_rate"] = (learning_rate)
  params["iterations"] = int(iterations)
  params["l2_leaf_reg"] = (l2_leaf_reg)
  params["random_strength"] = (random_strength)
  params["border_count"] = int(border_count)
  params["colsample_bylevel"] = (colsample_bylevel)
  params["leaf_estimation_iterations"] =
int(leaf_estimation_iterations)
  scores = cgb.cv(cv_dataset,
        params,
        fold_count=5, num_boost_round=1000, early_stopping_rounds=10)
  return np.max(scores['test-RMSE-mean'])
```

```
# Set hyperparamter bounds

pds = {
  'depth': (1, 16),
  'bagging_temperature': (0.0001, 10),
  'learning_rate': (0.0001, 1),
  'iterations': (1, 50000),
  'l2_leaf_reg': (0.0001, 10),
  'random_strength': (0.0001, 10),
  'border_count': (1, 255),
```

```
  'colsample_bylevel': (0.0001, 1),
  'leaf_estimation_iterations': (1, 10)
  }


# Surrogate model

optimizer = BayesianOptimization(hyp_cat, pds, verbose=2,
random_state=42)


# Run optimiser

start_time = time.time()
optimizer.maximize(init_points=10, n_iter=50)
print("--- %s seconds ---" % (time.time() - start_time))


# The best combination of parameters and target value found can be
accessed via the property optimizer.max.

optimizer.max


# While the list of all parameters probed and their corresponding
target values is available via the property optimizer.res.
for i, res in enumerate(optimizer.res):
  print("Iteration {}: \n\t{}".format(i, res))
```

APPENDIX 7

Python Algorithm for PDP (sklearn.inspection module)

```
# PDP
common_params = {
  "subsample": 50,
  "n_jobs": 1,
  "grid_resolution": 200,
  "random_state": 42
}
print("Computing partial dependence plots...")
features_info = {
  # features of interest
  "features": ['Elevation (m)','Slope (°)','Aspect (°)','Surface
Roughness','Topographic Position Index','Terrain Ruggedness
Index','Terrain Surface Texture','Vector Ruggedness Measure'],
  # type of partial dependence plot
  "kind": "average"
}
_, ax = plt.subplots(ncols=4, nrows=2, figsize=(10, 6),
constrained_layout=True)
display = PartialDependenceDisplay.from_estimator(
  hgb_model,
  X,
  **features_info,
  ax=ax,
  **common_params,
)
plt.show()
```

APPENDIX 8

Python Algorithm for PDP (PDPbox package)

```
# PDP Plot - Slope

feature_to_plot = 'Slope (°)'
pdp_slope = pdp.pdp_isolate(model= cat_model_best, dataset=X_test,
model_features=features, feature=feature_to_plot)
plot_params = {
  # plot title and subtitle
  'title': 'Slope (°)',
  'title_fontsize': 12,
  'subtitle_fontsize': 12,
  'font_family': 'Arial',
  # matplotlib color map for ICE lines
  'line_cmap': 'Blues',
  'xticks_rotation': 0,
  # pdp line color, highlight color and line width
  'pdp_color': '#1A4E5D',
  'pdp_hl_color': '#FEDC00',
  'pdp_linewidth': 2,
  # horizon zero line color and with
  'zero_color': '#E75438',
  'zero_linewidth': 2,
  # pdp std fill color and alpha
  'fill_color': '#66C2D7',
  'fill_alpha': 0.2,
  # marker size for pdp line
  'markersize': 5,
}
pdp.pdp_plot(pdp_slope, feature_to_plot, figsize=(8,8),
plot_pts_dist=False, center=True, plot_lines=False,
plot_params=plot_params)
plt.show()
```

APPENDIX 9

Python Algorithm for SHAP

```
# SHAP VALUES
import shap


# Fits the explainer

start_time = time.time()
explainer = shap.Explainer(cat_model_best, seed=42)
print("--- %s seconds ---" % (time.time() - start_time))


# Calculates the SHAP values.

start_time = time.time()
shap_values = explainer(X_test)
print("--- %s seconds ---" % (time.time() - start_time))


# Plots

shap.summary_plot(shap_values, X_test, plot_type="bar")
shap.summary_plot(shap_values)
shap.plots.bar(shap_values)
shap.summary_plot(shap_values, plot_type='violin')
```