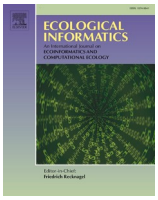
Ecological Informatics 73 (2023) 101934

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Hierarchical Bayesian geostatistics for C stock prediction in disturbed plantation forest in Zimbabwe

Tsikai S. Chinembiri [a,*], Onisimo Mutanga [a], Timothy Dube [b]

[a] University of KwaZulu-Natal, College of Agricultural, Earth and Environmental Sciences, P Bag X01 Scottsville, 3209 Pietermaritzburg, South Africa
[b] University of the Western Cape, Department of Earth Sciences, Private Bag X17, Bellville 7535, South Africa

ARTICLE INFO

Keywords:
Geostatistics
Bayesian inference
Markov Chain Monte Carlo
Hierarchical modelling
C stock
Bayesian prediction
Mean squared shortest distance

ABSTRACT

We develop and present a novel Bayesian hierarchical geostatistical model for the prediction of plantation forest carbon stock (C stock) in the eastern highlands of Zimbabwe using multispectral Landsat-8 and Sentinel-2 remotely sensed data. Specifically, we adopt a Bayesian hierarchical methodology encompassing a model-based inferential framework making use of efficient Markov Chain Monte Carlo (MCMC) techniques for assessing model input parameters. Our proposed hierarchical modelling framework evaluates the influence of two but related covariate information sources in C stock prediction in order to build sustainable capacity on carbon reporting and monitoring. The perceived improvements in the spectral and spatial properties of Landsat-8 and Sentinel-2 data and their potential to predict C stock with shorter uncertainty bounds is tested in the developed hierarchical Bayesian models. We utilized the Mean Squared Shortest Distance (MSSD) as the objective function for optimization of sampling locations for equal area coverage. Specifically, we evaluated the models using four selected remotely sensed vegetation indices namely, the normalised difference vegetation index (*NDVI*), soil adjusted vegetation index (*SAVI*), enhanced vegetation index (*EVI*) and an additional distance to settlements anthropogenic variable that justifies from the history of the studied plantation forest in the eastern highlands of Zimbabwe. We evaluated two models making use of Landsat-8 and Sentinel-2 derived predictors using the Root Mean Squared Error (*RMSE*), Mean Absolute Error (*MAE*), Coverage (*CVG*) and Deviance Information Criteria (DIC). The Sentinel-2 based C stock model resulted in RMSE of 1.16 $MgCha^{-1}$, *MAE* of 1.11 $MgCha^{-1}$, *CVG* of 94.7% and a *DIC* of −554.7 whilst its Landsat-8 based C stock counterpart yielded a *RMSE*, *MAE*, *CVG* and *DIC* of 2.69 $MgCha^{-1}$, 1.77 $MgCha^{-1}$, 85.4% and 43.1 respectively. Although predictive models from both sensors show great improvement in predictive accuracy when modelling the spatial random effects, the Sentinel-2 based C stock predictive model substantially outperforms its Landsat-8 based C stock counterpart. The Sentinel-2 based C stock predictive hierarchical model therefore adequately addresses multiple sources of uncertainty inherent in the spatial prediction of C stock in disturbed plantation ecosystems. It is evident from the results of this study that carbon reporting and monitoring can always be improved by scouting for improved and easily accessible remote sensing data and allow forest practitioners to keep track of error across space in resource environments of interest.

1. Introduction

The importance of forests as carbon sinks and the necessity to preserve, monitor and enhance terrestrial carbon stocks is recognised in the Kyoto Protocol of the United Nations Framework Convention on Climate Change (UNFCCC). This is because changes in forest carbon stocks influence the atmospheric carbon dioxide (CO_2) concentration (Millington & Townsend, 1989; Fan et al., 2022). Quantifying the uncertainty associated with forest carbon stock estimation and prediction can be enhanced by the inclusion of historical information such as scales of spatial variability known for characterising carbon stock dynamics in particular forest ecosystems (Do et al., 2022; González-Vélez et al., 2021). Surveys involving the collection of biomass data from forest plantations are time consuming and expensive. Yet, players and practitioners in the timber industry can always capitalise on historical data and expert opinion to quickly save as aids for accurate accounting and monitoring of carbon stock at landscape scales.

Natural forests in the sub-Saharan African region face increasing

* Corresponding author.
E-mail addresses: chinembiri24500@alumni.itc.nl (T.S. Chinembiri), MutangaO@ukzn.ac.za (O. Mutanga), tidube@uwc.ac.za (T. Dube).

https://doi.org/10.1016/j.ecoinf.2022.101934
Received 10 May 2022; Received in revised form 25 November 2022; Accepted 25 November 2022
Available online 28 November 2022

threats from cultivation, grazing and urban growth (Traore and Tiegu-hong, 2018). The rising demand of wood for both industrial and household energy and, in recent times, for carbon sequestration, rationalises the adoption of plantations as a viable and sustainable option for meeting these demands. For example, Zimbabwe lost approximately $3 billion in potential revenue and more than 4000 jobs due to deforestation by settlers at the height of the Fast Track Land Reform Programme (FTLRP), twenty-two years ago (Newsday, 2017). Reports from the Timber Producers Federation (TPF) published in 2014 indicated that the country's timber plantations were nearly facing total collapse with the national timber industry declining by 25% (Shumba & Marongwe, 2016).

An analysis of a mangrove forest in Vietnam making use of remote sensing data and Artificial Neural Network (ANN) gave mangrove Above Ground Biomass (AGB) predictions which ranged from 6.53 to 368.2 $Mgha^{-1}$ in 2000 and from 13.75 to 320.3 $Mgha^{-1}$ in 2020, respectively (Do et al., 2022). AGB predictions made by Fararoda et al. (2021) using machine learning methods including AdaBoost, random decision forest, multilayer neural networks and Bayesian ridge regression endorsed random forest and AdaBoost as the best performing methods. It is acknowledged in literature that the type of sensor and the method of prediction influence the performance of AGB models. For instance, Fassnacht et al. (2014) compared LiDAR and hyperspectral data for modelling AGB and concluded that LiDAR data fused with Random Forest (RF) models offered the best AGB predictive model.

Zimbabwe has a long-established commitment to conservation of biodiversity and its sustainable use as 16% of the country's land was under reserved forest and national parks before the 2000 Fast Track Land Reform Programme (FTLRP) (FAO, 2018). Subsequently, the country signed and ratified the Convention on Biological Diversity (CBD) as a recognition of the role of natural resources to the national economy. Most of the forest disturbances being experienced in the forest plantations are blamed on the government's lack of support for the resettled farmers in the post Agrarian Land Reform period. This has resulted in increased interface between communities and forest areas formerly designated for timber plantations. In fact, the amount of additional biomass that can be accumulated in these areas depend much on the forest condition and management practices in place (Leenhouts, 1998).

Obtaining useful estimates of uncertainty relating to forest carbon stock with imprecise predictions needed for decision making using model-based geostatistics is an established problem in climate change and carbon inventory studies (Ravindranath and Ostwald, 2008). Design-based estimation methods assume error propagation from sampling design and can therefore be appropriately accounted for if sampling plots are probabilistically selected (Cochran, 1977; Thompson, 2002). On the other hand, errors are attributed to the underlying process by which the outcome variable, that is, carbon stock, is generated in model-based assessments (Gregoire, 1998; Ver Hoff, 2002). A well-known limitation of error maps associated with classical geostatistical estimation techniques like co-kriging and kriging is their inability to account for uncertainty in the variogram-derived spatial covariance parameters (Cressie, 1993). Non-hierarchical implementations of spatial predictions struggle to efficiently deal with uncertainty associated with spatial covariance parameters, that is, spatial decays and spatial variances (Diggle and Ribeiro Jr, 2007). The advantages of choosing a Bayesian hierarchical methodology to inference over other related techniques include access to the entire posterior predictive distribution (PPD) (Goulard and Voltz, 1992). According to Beloconi and Vounatsou (2020), full access to the PPD facilitates subsequent analysis that informs management and ecological objectives whilst accounting for prediction uncertainty.

Hierarchically constructed Bayesian geostatistical models are a potential remedy to the challenges faced when utilizing the aforementioned approaches (Gelfand et al., 2004). Incorporation of medium and high spatial resolution remote sensing derived auxiliary information measured at the same geo-location as the independent variable further enhances the ability of hierarchical models in providing resource estimates with a reduced measure of uncertainty. Babcock et al. (2015) utilized LIDAR for predicting forest biomass using the hierachical Bayesian approach and established additional predictive performace after accounting for spatial random effects to the prediction equation. Significant progress has also been made in mapping forest attributes at broader spatial scales using spatially enabled hierarchical models. Such forest attributes include forest biomass (Finley et al., 2011; Johnson et al., 2014), community forest composition (Finley et al., 2009), rate of deforestation (Agarwal et al., 2005) and specific tree structure attributes (Babcock et al., 2012). Making predictions of forest biomass and other forest parameters needed for carbon accounting under UNFCC using Bayesian geostatistics comprehensively takes parameter uncertainty into account. However, despite the improvements coming from new generation remote sensing platforms as the Landsat series and the European Space Agency (ESA) based multispectral Sentinel products, no study has utilized vegetation indices from new generation sensors, particularly Landsat-8 and Sentinel-2, as predictors of C stock under a Bayesian Hierarchical framework in climate change studies.

An understanding of the carbon cycle and reduction in carbon emissions can be done through the global monitoring of AGB forest. Nevertheless, there is still a lot of uncertainty regarding the quantity and spatial distribution of AGB as a result of difficulties faced when measuring AGB using field measurement standards (Lefsky, 2010; Simard et al., 2011). Space-borne remote sensing techniques can collect data correlated with AGB spatial distribution over large national and global regions in a cost-effective method (Clerici et al., 2016). The limited sensitivity of earth observation sensors to AGB and the lack of in situ data needed for calibration at the appropriate scales useful for remote sensing are some of the challenges of facing these C mapping methods. Obtaining correct allometric models in certain regions of the world is a problematic because of armed conflicts, remoteness of the areas of interest and lack of capacity to generate plausible data. Remote sensing signal saturation due to high density of AGB, clod cover, especially in tropical regions and complexity of signal retrieval due to complicated topography are also some of the major AGB mapping challenges (Song et al., 2010).

Semela et al. (2020) recommend both Sentinel-2 and Landsat-8 as valid information sources for grass biomass estimation in mountainous environments using random forest modelling. Astola et al. (2019) compared Sentinel-2 and Landsat-8 models for retrieving leaf area index (LAI), canopy cover (CC) and effective canopy cover (ECC) using spectral bands available for both sensors and found non-systematic differences between the Sentinel-2 and Landsat-8 data. Korhonen et al. (2017) also compared Sentinel-2 and Landsat-8 imagery for forest variable prediction and established Sentinel-2 to outperform Landsat-8 due to the enhanced spatial resolution in the former compared to the later. The majority of studies comparing Sentinel-2 and Landsat-8 for predicting AGB favour Sentinel-2 over Landsat-8. This is further justified from the work of Jha et al. (2021) who compared Worldview-3, Sentinel-2 and Landsat-8 for mapping AGB in a forest landscape in Thailand and established Worldview-3 and Sentinel-2 as better predictors than Landsat-8 owing to the red-edge and the higher spatial and spectral properties of Worldview-3 and Sentinel-2. It is therefore perceived that the improvements in the spectral and spatial properties of Sentinel-2 over Landsat-8 can reduce prediction uncertainty on C stock prediction under a fully modelled Bayesian Hierarchical framework.

The purpose of this paper is therefore to develop and evaluate the performance of Bayesian hierarchical models employing Landsat-8 Operational Land Imager (OLI) and Sentinel-2 derived vegetation indices for carbon stock prediction in disturbed plantation ecosystems in Zimbabwe. The employment of the Bayesian methodology coupled with predictor information from new generation remote sensing predictors is expected to offer better C stock predictions in C stock monitoring an reporting. Model evaluation is utilized to identify the strengths and

deficiencies of candidate models and hence, the trade-off associated with utilization of different information sources in C stock prediction. We further assess and compare C stock prediction uncertainty intervals generated from Landsat-8 OLI and Sentinel-2 using the Bayesian hierarchical framework. This is achieved through the development of separate but related models, integrating Landsat-8 OLI and Sentinel-2 derived spectral indices. These models are constructed using a Bayesian hierarchical geostatistical framework employed for predicting C stock in disturbed plantation forest in Zimbabwe. The paper aims at providing clinical information needed for decision making by governments and Non-Governmental Organizations (NGOs) working towards the fulfilment of the 2030 Agenda on Sustainable Development and the Paris Agreement on climate change. The government's pursuit for charming private capital into the timber industry in recent years in Zimbabwe can only be taken seriously if C stock estimates in the targeted plantation ecosystems for reforestation and afforestation programmes are provided with an accompanying measure of uncertainty.

2. Methods

2.1. Study area

The study was carried out at lot 75 A of Nyanga Downs in Nyanga district in the Eastern Highlands of Zimbabwe, (Fig. 1). The study area is dominated by *Eucalyptus grandis*, *Eucalyptus camaldulensis* and *Pinus patula* plantation forest species which has some of its patches being used for agriculture, grazing and gold panning and is located between latitude 32°40' E and 32°54' E and 18°10'12"S and 18°25'4"S longitude as illustrated in Fig. 1. Grazing, agriculture and gold panning activities came after part of the commercially owned plantation forests were redistributed to small and medium sized indigenous farmers in 2000. This development has increased interface between settlements and timber plantations in all forests originally designated under forest plantations in Zimbabwe.

The study domain covers an area of approximately 2767 ha. Rainfall amounts are variable and range from 741 mm to 2997 mm with a mean annual precipitation of 1200 mm. Annual mean temperatures range from a minimum of 9 °C to 12 °C to a maximum of 25 °C to 28 °C. The weather is very hot and extensive wild fires occur in the high-elevation grasslands from August to November where the grasses are dry.

2.2. Remote sensing derived predictors

The spatial coverage and extent of RS data such as Landsat in addition to their relatively low cost vindicate their usefulness as covariates. Sentinel-2 comprises six land monitoring bands which are comparable to Landsat-8 and also includes three additional bands encompassing the red-edge (RE) spectrum (Drusch et al., 2012). The Red edge bands are cantered at 704, 740, and 782 nm with bandwidths of 15, 15, and 20 nm, respectively. The red-edge is the major spectral feature of vegetation positioned between the red absorption maximum (680 nm) and the high reflectance in the Near-Infrared (750 nm) (Frampton et al., 2013). In addition, Sentinel-2's land surface bands have a spatial resolution of 10 m and 20 m in comparison to Landsat-8's 30 m bands. These sensor

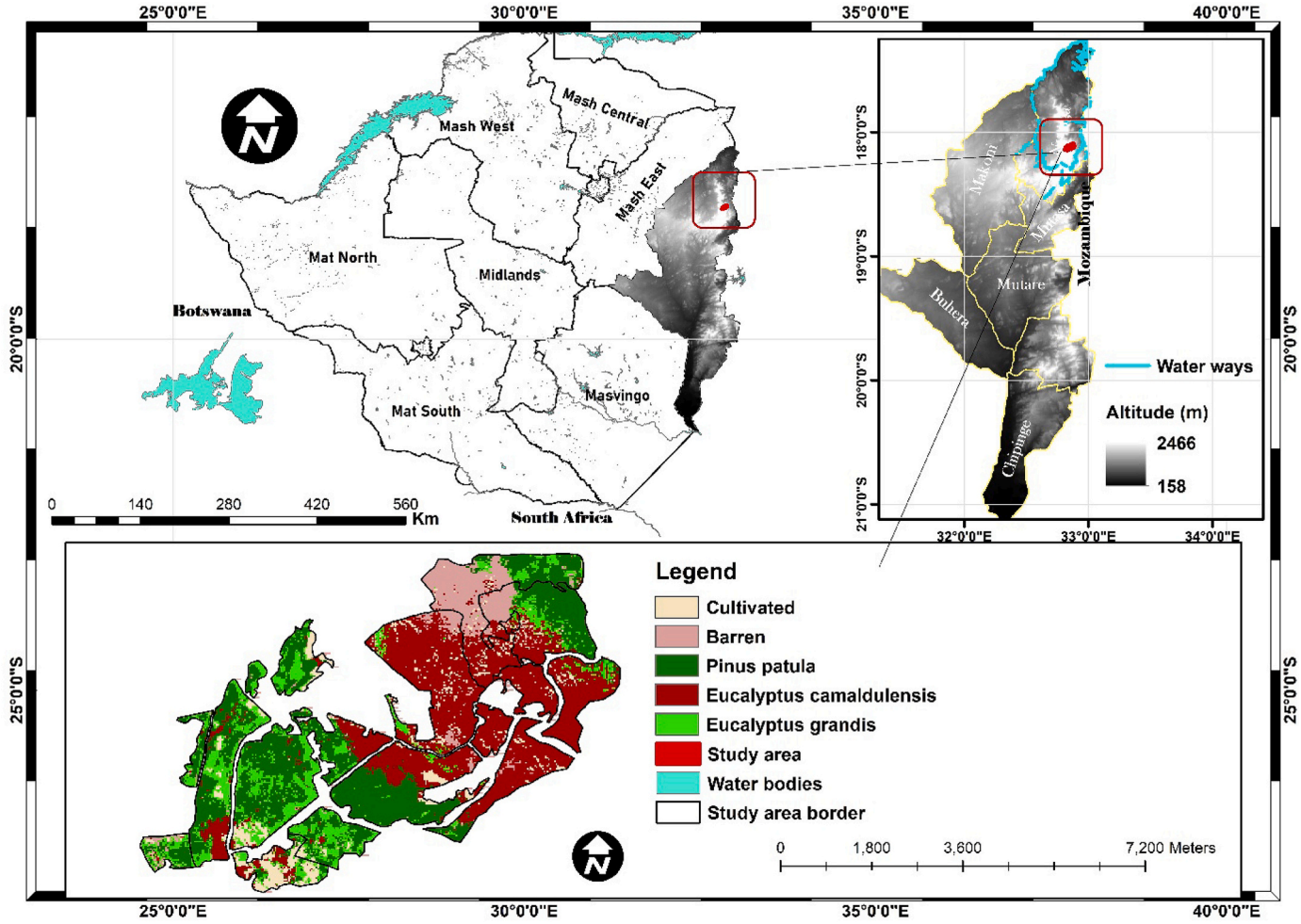


Fig. 1. Map of the study area indicating (a) the province where samples were derived, (b) study area location within the particular province and (c) the spatial distribution of the sampled regionalised variable in the lower panel.

differences provide the basis and an inferential framework on which to learn and understand the sensor that outperforms the other in C stock prediction.

2.2.1. Landsat OLI and Sentinel-2 MSI imagery

Landsat-8 imagery was obtained from the United States Geological Survey Earth Explorer (http://earthexplorer.usgs.gov) as analysis-ready datasets (ARDs). Datasets were filtered with cloud cover and cloud shadow cover thresholds set to less than 10%. Sentinel-2 cloud-free imagery was acquired on 20 September 2020 at the same time as the Landsat-8 OLI data collection covering the entire area of interest at lot 75A Nyanga Downs in the eastern highlands of Zimbabwe.

Sentinel-2 imagery is taken using the multispectral instrument (MSI), a push-broom imaging instrument that measures the Earth's top of atmosphere (TOA) reflected radiance in thirteen (13) spectral bands ranging from 443 nm to 2190 nm. The Sentinel-2 data were derived as level-1C 12-bit pre-set TOA reflectance values. The pre-processing and orthorectification of the level 1-C products was carried out in the R statistical and computing environment using the *sen2r* package (Ranghetti et al., 2020).

We utilized Normalised Difference Vegetation Index (*NDVI*), Soil Adjusted Vegetation Index (*SAVI*), Enhanced Vegetation Index (*EVI*) and an additional distance to settlements anthropogenic variable as covariates for C stock prediction in disturbed plantation forest. The aforementioned vegetation indices have been applied in literature, including, Li and Li (2019) and Bordoloi et al. (2022), as independent variables in AGB biomass estimation. Previous research on climate studies including Wang et al. (2005) and Cross et al. (2010) have modelled the effects of anthropogenic or biophysical variables on biomass separately. In addition, a limited number of variables for each of these classes of predictors have been applied in biomass estimation. The present study employs a combination of these classes of independent variables using a Bayesian hierarchical approach for C stock estimation and prediction.

2.3. Sampling design

2.3.1. Spatial coverage sampling scheme

We utilized the Mean Squared Shortest Distance (MSSD) as the objective function for optimization of sample locations in the current study. We therefore employed the *k*-means clustering algorithm for equal area coverage sampling. Walvoort et al. (2010) demonstrated how the mapping of regionalized variables can be enhanced by uniform dispersal of sampling locations within a study domain. Furthermore, Brus et al. (2006) showed how uniform coverage of the study domain with sampling locations can both be used for resolving the mapping and estimation of spatial means of regionalized variables in the soil, forestry and environmental research.

Amongst the methods that have been utilized for optimization of the sampling pattern in literature, including, Spatial Simulated Annealing (SSA), the MSSD stands out as the most suited optimization methodology for both prediction and estimation of regionalized variables. The design is especially appropriate in scenarios where sampling schemes cannot be extended beyond a single phase and where the area has never been sampled before.

Fig. 2. Spatial coverage sampling design.

2.3.2. Geographical partitioning of the study domain by the k-means algorithm

The study area was subdivided into compact subregions by clustering of the raster cells making up the study region using the k-means optimization algorithm (Brus et al., 2006, Walvoort et al., 2010).

The *k*-means optimization algorithm makes use of the *x* and *y* coordinates of the centre points of the raster cells as classification variables. Centroids of the clusters were then used as sample points where sampling plots for C stock were established as Illustrated in Fig. 2.

2.4. Carbon stock data

2.4.1. Above ground tree biomass (AGTB) field measurement

We sampled and took measurements of all trees with at least 10 cm diameter at breast height (at 1.3 m above the soil surface) using 500 m^2 circular supports using diameter and linear tapes from the 19th of September 2021 to the 24th of October 2021. According to Gibbs et al. (2007), trees with DBH less than 10 cm have insignificant C stocks. Slopes in the study area are generally below 30% and hence, slope correction was not considered during the tree biomass measurements (Ravindranath and Ostwald, 2008). Optimization of the sample locations that resulted in 200 spatial coverage samples was carried out using the spcosa-package in the R Statistical and Computing Environment (Brus et al., 2006, Walvoort et al., 2010). The 200 sampling locations were pre-loaded into a 72H handheld Garmin GPS before launching the field work exercise. However, 191 observations of forest biomass were obtained during the sampling phase as nine sampling plots fell outside the boundaries of the defined study area (Fig. 2). Statistics of sampled C stock data for the measured plantation forest parameters from the field for the studied region are illustrated in Table 1.

2.4.2. Biomass calculation and derivation of C stock

Pinus species biomass was calculated using allometric equations from Brown (1997) whilst biomass of the Eucalyptus species was calculated using allometric equations proposed by Zunguze (2012). Tree diameter was measured at 1.3 m above the soil surface with all tree species with at least 10 cm measured. Application of the aforementioned allometric equations was befitting as the same equations were applied to Eucalyptus and Pinus species in Manica province in Mozambique which closely resembles and approximates the weather and climatic conditions of the eastern highlands of Zimbabwe. The above ground biomass of every individual species was then converted to C stocks per species by means of a conversion factor of the IPCC (2006). Per plot (support) estimated values were then expanded to a standard unit area, in this case, a hectare ($MgCha^{-1}$).

We took the position that the relationship between AGB and remote sensing features apart from being determined by the crown areas of the measured and sampled plantation species in Fig. 1, but also determined by other factors like tree age, planting density and plantation conditions (i.e. climate or soils). We therefore assumed that the AGB model captures all these factors indirectly since the remote sensing signal is also a product of these factors.

2.5. Hierarchical Bayesian modelling

We adopted the Bayesian hierarchical framework in order to fully account for parameter uncertainty in the sampled C stock. The Bayesian hierarchical framework consists of four stages as detailed below:

2.5.1. Observed C stock data

Let $Y(s)$ denotes the observed log transformed C stock data at a spatial location s, $s = 1, \ldots, S$, then,

$$Y(s) = X^T(s)\beta + w(s) + \varepsilon(s) \quad (1)$$

where

$w(s)$ is a spatial random effect term representing the effect of being at location s, and in our case, an exponential autocorrelation function.

β is a vector of regression coefficients associated with a $1x\ q$ vector of predictors $X(s)$ sampled at location s.

$\varepsilon(s)$ represents the random error (white noise) assumed *i. i. d.* $N(0, \sigma_\varepsilon^2)$

2.5.2. Spatial random effects specification

We assumed the spatial random effects $w(s) = (w, \ldots, w_s)^T$ arise from a multivariate distribution, that is, a Gaussian random field, scaled by a spatial variance σ_w^2 and correlations proportional to the separation distance, d_{ij}, between sites as follows:

$$w \sim MVN\left(0_s, \sigma_w^2 \Sigma_w\right) \quad (2)$$

Where 0_s is an $S\ x\ 1$ vector of zeros, σ_w^2 representing the between site variance, whilst Σ_w is the $S\ x\ S$ correlation matrix with elements (s_i, s_j) indicating the correlation between sites s_i and s_j, $i, j = 1, \ldots, S$. We assumed a stationary model with an isotropic covariance structure where the correlation between sites s_i and s_j is a function of the separation distance between the sites. We therefore computed the covariance matrix Σ_w for the correlation function as follows:

$$f(ds_i s, \phi, \kappa) = exp\left(-\phi d^{\kappa} s_i s_j\right)$$

Because of limited information in the data to be able to estimate both κ and ϕ, we followed common practice and fixed κ to be one (Diggle and Ribeiro Jr, 2007).

2.5.3. Model prediction

A fully Bayesian hierarchical framework entails simultaneous estimation and prediction of model parameters. Uncertainty associated with estimates of model coefficients are considered and supplied through the model to the predictions. The Markov Chain Monte Carlo algorithm was used to make and estimate predictions at a location s' with the form:

$$\widehat{Y}(s)^{'} = X^T(s)^{'}\widehat{\beta} + \widehat{w}(s)^{'} \quad (3)$$

As such, predictions were made using a combination of the overall mean, the spatial effect and the effect of independent variables from medium resolution sensor satellite data. We calculated the spatial component of the predictions using multivariate normal properties. That is, if $\boldsymbol{w} = (w_1, \ldots, w_s)'$ are the spatial effects at site of C stock observation, then the conditional distribution at an arbitrary location, s′, is $w_{s'} \mid \boldsymbol{w}$ that is normally distributed with mean and variance denoted by:

Table 1
Summary statistics of the measured C stock plantation forest parameters.

Statistic	*Eucalyptus camaldulensis*			*Eucalyptus grandis*			*Pinus patula*		
	DBH	Height	C stock	DBH	Height	C stock	DBH	Height	C stock
Mean	81.4	60.6	2485.3	67.4	70.6	405.7	56.8	58.6	377.9
Median	77.4	52.7	1470.3	51.4	49.7	327.8	43.5	38.7	295.4
Max	231.9	88.9	8998.2	97.9	90.1	429.8	64.3	66.6	600.3
Min	11.4	23.8	13.7	14.7	27.8	111.3	10.6	19.4	9.7

$$E[w_{s'}|\boldsymbol{w}, Y] = \sigma_w^{-2}\delta_j'\Sigma_w^{-1}\boldsymbol{w} \quad (4)$$

And

$$var(w_{s'}|\boldsymbol{w}, Y) = \sigma_w^2\left(1 - \delta_j'\Sigma_w^{-1}\delta_j\right) \quad (5)$$

respectively, where δ_j is the vector of the effect of distance between new sites and C stock observation locations where each element $\delta_{ij} = f(d_{s_i s_j}, \phi)$.

2.5.4. Hyperprior specification

We performed Bayesian hierarchical modelling using the *spBayes* package (Finley et al., 2007) in the R Statistical and Computing environment (R Core Development Team, 2008). As articulated in Gelfand (2012), the Bayesian approach treats the vector of model parameters, $\boldsymbol{\theta} = \boldsymbol{\beta}, \sigma^2, \phi, \tau^2$ as random and mutually independent variables and assigns prior distributions to the parameters. By taking $p(\boldsymbol{\theta})$ to be a prior distribution and $p(\boldsymbol{\theta}|y, X)$ to be the posterior distribution of the observed C stock model parameters, we computed the posterior distribution of model parameters, $\boldsymbol{\theta}$, as follows:

$$p(\boldsymbol{\theta}|\boldsymbol{y}, X) \propto p(\boldsymbol{\theta})\ x\ N(\boldsymbol{w}|0, \Sigma_w)\ x\ N(\boldsymbol{y}|X'\boldsymbol{\beta} + \boldsymbol{w}, \Sigma_\varepsilon) \quad (6)$$

A summarised modelling framework for the hierarchical specification shown from Eq. (1) to Eq. (7) is illustrated in the flow chart illustrated in Fig. 3. We utilized Eq. (6) to quantify uncertainties in model parameters. Predicted C stock values were then derived from the predictive distribution of C stock at unsampled locations, s_0 using Eq. (7), which quantifies uncertainties in C stock predicted values.

$$p(y_0|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) \propto \int p(y_0|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{x}_0)p(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{X})d\boldsymbol{\theta} \quad (7)$$

Where; y_0 denotes the predicted C stock at a location s_0 and x_0 are the covariate values at location s_0.

We assigned normal prior for the overall mean of measured C stock and a multivariate normal prior for the regression coefficients. We also classified model parameters into two categories of (1) regression coefficients consisting of $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)'$ in the mean function and (2) the partial sill (σ^2), nugget (τ^2) and range (ϕ) in the variance covariance matrix. In respect of our sampled data, we specified two sets of priors for hierarchical modelling of C stock using predictors from Landsat-8 and Sentinel-2 data. An inverse gamma prior distribution was assigned for the data and measurement error variance whilst a uniform prior was assigned for the spatial decay parameter, ϕ as defined by $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$ below:

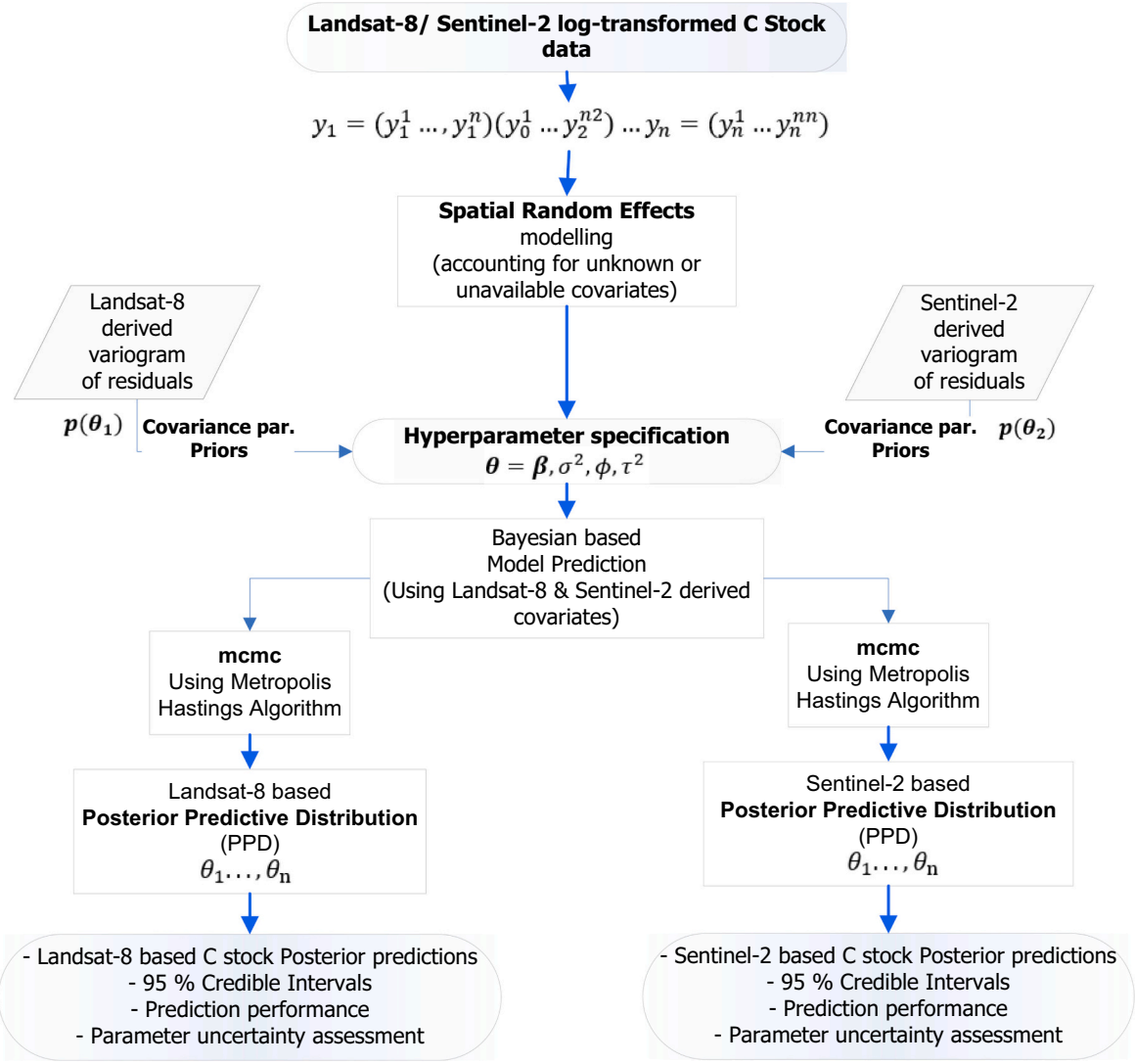


Fig. 3. Hierarchical Bayesian geostatistical modelling illustrated by Eqns. 1–7 for C stock prediction.

$$p(\boldsymbol{\theta}_1) = \boldsymbol{Unif}(\phi|0.38, 0.0012)\ x\ \boldsymbol{IG}\left(\sigma^2|0.75, 1.76\right)\ x\ \boldsymbol{IG}\left(\tau^2|0.1, 1.76\right)\ x\ \boldsymbol{MVN}\left(\beta|\mathbf{0}, \Sigma_\beta\right)$$

$$p(\boldsymbol{\theta}_2) = \boldsymbol{Unif}(\phi|0.38, 0.0012)\ x\ \mathbf{IG}\left(\sigma^2|0.071, 0.021\right)\ x\ \boldsymbol{IG}\left(\tau^2|0.071, 0.0028\right)\ x\ \boldsymbol{MVN}\left(\beta|\mathbf{0}, \Sigma_\beta\right)$$

Where;

$p(\boldsymbol{\theta}_1)$ and $p(\boldsymbol{\theta}_2)$ are prior specifications for Landsat-8 and Sentinel-2 derived covariates, respectively.

In both sets of priors ($\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$), the shape parameters were chosen in order to guarantee that the prior distribution has finite mean and the distribution is adequately vague that it would not have finite variance (Demirhan and Kalaylioglu, 2015). In the same vein, we adopted scale parameter values in order to expresses the preference that the prior mean of σ_ε^2 is less than that of $\sigma_{\boldsymbol{w}}^2$. This is vindicated from the fact that we expect the measurement error variance (nugget, σ_ε^2) to be much smaller than the data variance, $\sigma_{\boldsymbol{w}}^2$. The spatial decay parameter ϕ was assigned a uniform prior with support over the spatial range of the study domain.

In order to get results from the Bayesian hierarchical modelling, a Metropolis-Hastings algorithm typically used in the *spBayes* R package for Markov Chain Monte Carlo (MCMC) methods was employed (Gelman, 2006). An algorithm of one chain each with 20,000 MCMC iterations was specified for the posterior densities of the parameters in which 15,000 were discarded as burn-in.

2.6. Variogram modelling and exploratory analysis

We carried out spatial exploratory analysis of the sampled C stock observation data in order to get some insights regarding the choice of scale values on the prior distribution of the spatial covariance parameters. Spatial exploratory analysis is an important step towards assessing the strength of the spatial correlation structure of the modelled regionalized variable (Stoyan and Yaskov, 2014; Sahu, 2022; Pascual et al., 2022). We also streamlined the modelled C Stock data to the normality assumption through the Box-Cox transformation approach (Box and Cox, 1982). C Stock data was therefore transformed according to the Box and Cox (1982) formula illustrated in Eq. (8).

$$\begin{cases} Y(\lambda) = \begin{matrix} \left(Y^\lambda - 1\right)/\lambda \ if\ \lambda \neq 0 \\ log\ Y\ if\ \lambda = 0 \end{matrix} \end{cases} \tag{8}$$

Y is the C Stock observation and λ is the transformation parameter.

Developments made by Box and Cox (1964) regarding the transformation parameter in Eq. 8 are of the normal theory linear model as in Eq. (9),

$$y(\lambda) = \boldsymbol{X\beta}(\lambda) + \varepsilon \tag{9}$$

where $\boldsymbol{X}$ is an $n\ x\ p$ vector of predictors, $\boldsymbol{\beta}(\lambda)$ is a $p\ x\ 1$ vector of unknown parameters while the standard deviation of the independent errors $\varepsilon_i(i = 1, \ldots, n)$ is $\sigma(\lambda)$.

2.7. Model predictions and uncertainty assessment

We tested the statistical significance of the predictor variables by considering the 2.5% and the 97.5% percentiles of the posterior samples. Independent variables in the Landsat-8 and Sentinel-2 based hierarchical models were deemed statistically significant if their 95% Credible Intervals (CIs) excluded zero. Uncertainty of C stock predictions was assessed by employing the method of Hengl et al. (2004). Median predictions of C stock for each grid cell (10000m^2) were displayed using a colour ramp alongside the associated uncertainty (95% Credible Interval Widths (CIW)). Posterior predictive distribution CIWs was therefore utilized to assess the precision of model predictions. Highly uncertain predictions have wider 95% CIWs (Babcock et al., 2016). As documented in literature including Babcock et al. (2016) and Babcock et al. (2018), highly uncertain (less precise) predictions have wider CIW.

2.8. Model comparison and validation

Three different Bayesian hierarchical models constructed from Landsat-8 and Sentineal-2 derived predictors were compared as illustrated in Eq. (10), (11) and (12). The models compared were the independent error model (simple multiple linear regression), the spatial intercept only model and the spatial model, respectively

$$\boldsymbol{y} = \boldsymbol{x}\beta + \varepsilon \tag{10}$$

$$\boldsymbol{y}(s) = \boldsymbol{x}(s)\beta + \varepsilon(s) \tag{11}$$

$$\boldsymbol{y}(s) = \boldsymbol{x}(s)\beta + w(s) + \varepsilon(s) \tag{12}$$

Posterior samples of the spatial and spatial intercept only models' parameters and predictions are gathered via a MCMC algorithm and composition sampling. We further tested the predictive performance of each of the three candidate models using a k-fold cross validation algorithm, which proceeds by randomly splitting the 191C stock observations into ten almost equally seized segments (Duchene et al., 2016). Log-transformed C stock for the holdout data block was successively predicted given model parameters derived using data in the remaining nine blocks.

Root mean squared error (RMSE), Mean Absolute Error (MAE) and other validation statistics were calculated using the holdout posterior predicted means and observed C stock data for each of the three models (Green et al., 2020). We considered the model with the lowest k (k = 10)-fold RMSE and MAE as the "best" predicting model. Model convergence was assessed using graphical displays in the form of trace plots of the estimated model parameters (Jackman, 2000).

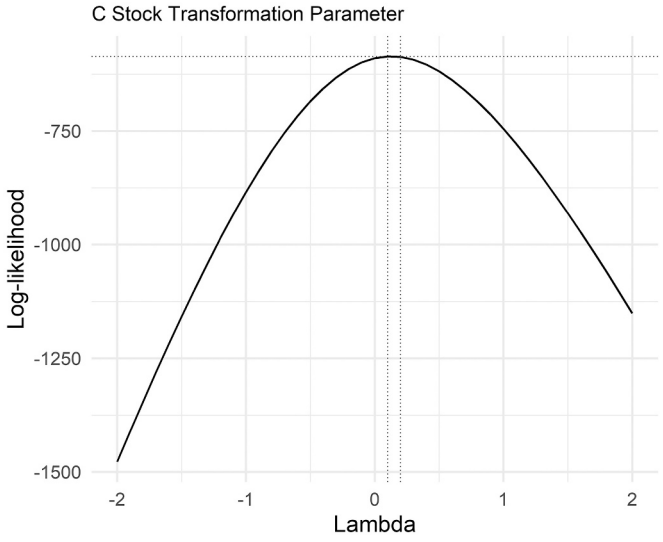


Fig. 4. Box-Cox Transformation for C stock.

3. Results

3.1. Exploratory analysis

An objective approach for arriving at the appropriate transformation parameter made through the Box-Cox transformation as in Eq. (8) for the C stock data led to a log-transformation as λ was estimated to be 0.012 as illustrated in Fig. 4.

A transformation parameter of 0.5 entails a square root transformation, a parameter of 1 entails no transformation whilst a transformation of 0 supports a logarithmic transformation as illustrated in Fig. 3 (Box and Cox, 1964).

As detailed in Section 2, parameter transformation from this preliminary analysis governed the treatment of the outcome variable in all the subsequent analysis involving treatment of the variable under different satellite derived auxiliary information. Log and square root transformation of outcome variables are common data transformation parameters in natural resources modelling (Babcock et al., 2016; Chinembiri et al., 2013).

The modelled C stock data employed in the Bayesian hierarchical framework displayed a positively skewed distribution as illustrated in Fig. 5. Positively skewed C stock data implies that the lower bounds of the sampled data are relatively lower than the rest of the data. A Box Cox transformation with a transformation parameter of 0.012 ($\lambda = 0.012$) entailed a log-transformation for the C Stock variable which led to an approximately normally distributed C Stock data as illustrated in Fig. 5 (b). The σ^2_{ϵ} and the σ^2_{w} variance hyperpriors for the Landsat-8 and Sentinel-2 based spatial models of the log-transformed C stock data were set according to the variogram of residuals of predictors derived from the respective medium resolution satellite sensors illustrated in Fig. 6(a) and Fig. 6(b) respectively.

As illustrated in Fig. 6, the σ^2_{ϵ} and the σ^2_{w} variance parameters of the C stock hierarchical Bayesian models derived from the respective medium resolution sensors show greater influence of Sentinel-2 derived predictors on C stock. Subsequent Bayesian hierarchical modelling of the response variable followed the trajectory of scale parameters illustrated in Fig. 6 as these provided a constraint on the prior probability distribution function of the parameter space.

3.2. Effects of Landsat-8 and Sentinel-2 predictors on C stock prediction

Covariate information in the form of *NDVI*, *SAVI*, *EVI* and distance to the nearest settlements from Landsat-8 and Sentinel-2 employed in the hierarchical modelling of C stock showed distance to nearest settlements and *NDVI* to be significant predictors of C stock. The two predictors are significantly different from zero as their 95% CIs exclude zero. The posterior distribution of predictor coefficients markedly differed between Landsat-8 and Sentinel-2 based spatial models. However, amongst the modelled predictors, *NDVI* and distance to settlements (*DIST*) were statistically significant in both the Landsat-8 and Sentinel-2 based spatial models as shown in Table 2.

The carbon stock model derived from Landsat-8 predictors implies a weaker spatial correlation with an effective range estimated as 1764 m and a 95% CI of (1578.9, 2307.7) whilst C stock model derived from Sentinel-2 independent variables shows a stronger correlation with an effective range of 2142.9 m with a 95% CI of (2000, 2307.7) meters directly estimated from the CI of the spatial decay parameter, ϕ. However, estimates of the range of spatial dependence from both medium resolution sensor C stock-based models seem plausible as the maximum distance between data locations within the study domain is 2463 m. As expected, we observe that the spatially structured variance, σ^2_{w}, is higher than the micro-scale variance in both models (Diggle and Ribeiro Jr, 2007).

*DIST** is the distance to the nearest settlements used in the Bayesian hierarchical modelling framework

3.3. C stock prediction from medium resolution sensor derived covariates

Fitted models with spectral independent variables from Landsat-8 and Sentinel-2 were employed in the prediction of C stock at unvisited locations in the study domain. Since spectral variables of *NDVI*, *SAVI*, *EVI* in addition to distance to settlements were on gridded raster with 10,000 m^2 resolution, the predicted log-transformed C stock represents averaged values in every pixel. MCMC sampling from the predictive distribution of the fitted models resulted in a total of 20,000 predicted samples for each grid. Consequently, predicted samples for each pixel provide a comprehensive account of the uncertainty in the predicted C stock values. The specific expressions and ratios of vegetation indices represent properties of green vegetation much better than individual bands (Baloloy et al., 2018).

As illustrated in Fig. 7, the 95% CIs for C stock predictions derived from Sentinel-2 spectral variables are much more plausible than predictors supplied from Landsat-8 spectral covariates. The indispensability of *NDVI* as an index well correlated with vegetation biophysical

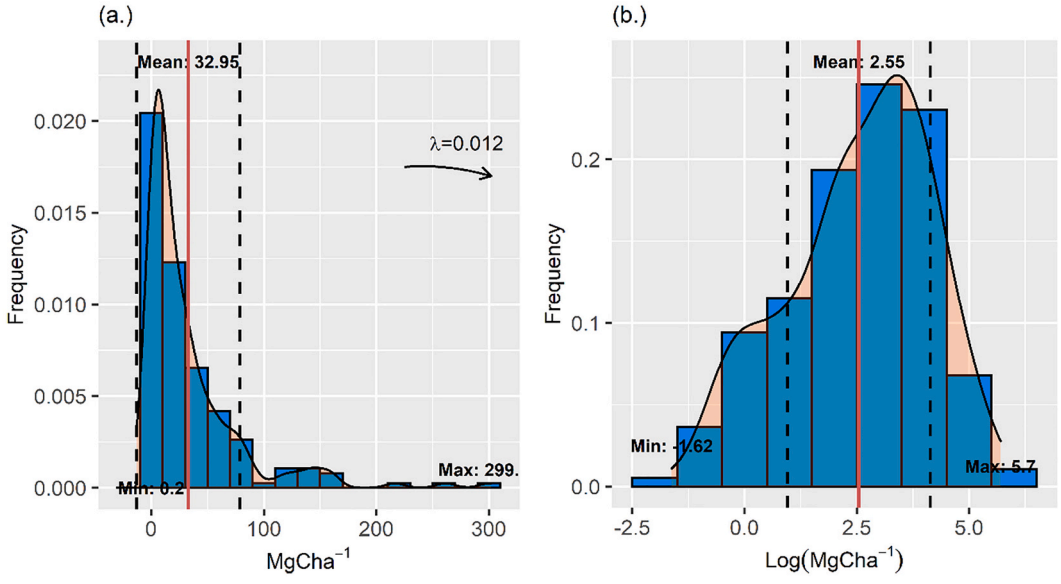


Fig. 5. Logarithmic transformation for the C stock data (with normal overlay).

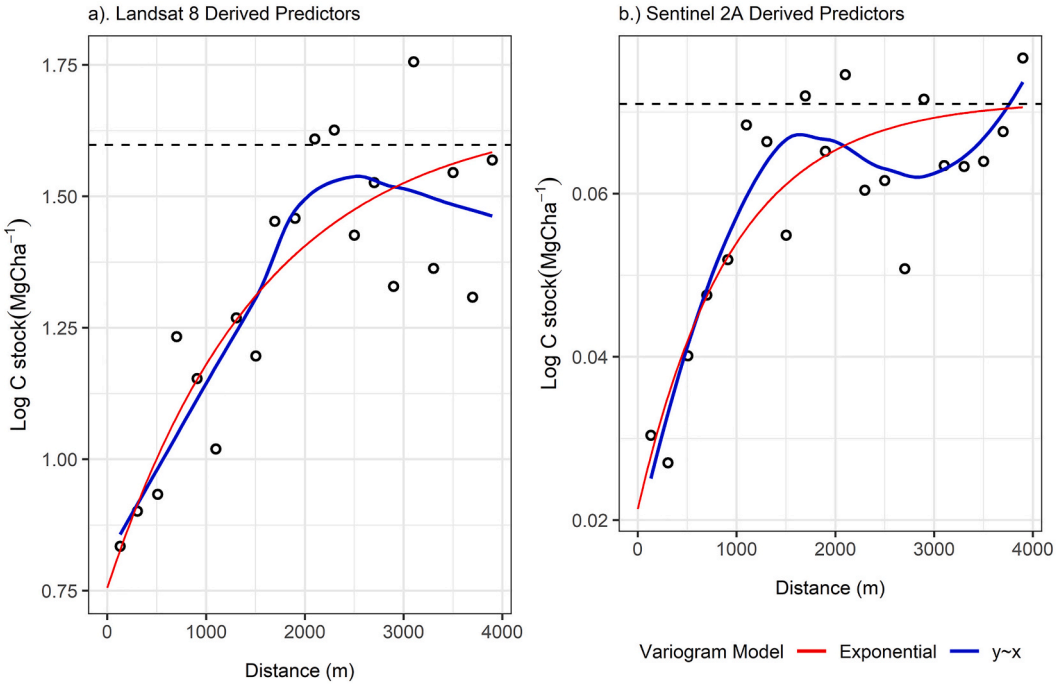


Fig. 6. Landsat-8 and Sentinel-2 derived variogram of residuals for the modelled C stock. The black dotted line is the asymptote of the theoretical variogram model.

properties, amongst them, leaf area index (*LAI*), and chlorophyll is much established (Baugh and Groeneveld, 2006). Sentiel-2 based C stock predictions have lower uncertainty compared to Landsat-8 OLI C stock aided predictions. This fact is supported by the evidence of the 95% posterior predictions illustrated in Fig. 8, showing C stock 95% CIs orders of magnitude lower than the predicted estimates in Sentinel-2 than those displayed in Landsat-8 based estimates. The accompanying credible intervals for both Landsat-8 and Sentinel-2C stock-based predictions are highly certain. This is because the entire part of the study area has 95% CIs orders of magnitude lower than the C stock estimates (Fig. 7 and Fig. 8).

It is observed that predictions in the central northern region of the studied region (Fig. 7) in Lot 75A of Nyanga Downs are much lower in Landsat-8 than in Sentinel-2 based C stock predictions. We attribute underprediction of C stock in this region in Landsat-8 based predictions to the finer spatial and spectral attributes in Sentinel-2 data in the visible and near-infrared portions of the electromagnetic spectrum (Sovdat et al., 2019; Wang et al., 2020; Ahmed et al., 2022). The enhanced spatial and spectral resolution of Sentinel-2 also justifies the shorter 95% CIW displayed by the Sentinel-2 based C stock predictions (1.02–1.82) MgCha^{-1} compared to Landsat-8 based C stock predictions (2.23–4.50) MgCha^{-1}. The results from both sensors look attractive compared to AGB reported in literature where Jiang et al. (2021) established an average AGB RMSE of 40.92 MgCha^{-1} in northeast China whilst Dang et al. (2019) reported a RMSE of 36.67 MgCha^{-1} in Vietnam using Random Forest (RF).

To add on, Takagi et al. (2015) utilized LiDAR for predicting forest biomass in Hokkaido, Japan, and established a biomass RMSE prediction of 19.10 MgCha^{-1}. Differences in the prediction accuracy between the reported results in literature and our study can also be explained by the differences in forest density as the aforementioned studies have been undertaken in tropical and subtropical rainforest biomes.

3.4. Performance assessment of C stock prediction models

Table 3 presents the *k*-fold cross validation statistics used to assess the predictive performance of Bayesian hierarchical models emanating from Landsat-8 and Sentinel-2 derived spectral predictor variables. Different predictive Bayesian hierarchical models from the two medium resolution satellite sensors, with varying levels of richness (independent error, spatial intercept only and the spatial models) favour the Sentinel-2 based C stock spatial predictive model (Table 3).

In addition to the application of RMSE, MAE and CRPS as model validation and performance criteria, coverage (CVG) is employed as an additional criterion of model performance. The Sentinel-2 based spatial C stock predictive model (Table 3) appears as the top performing model in terms of the RMSE (1.16 MgCha^{-1}), DIC (−554.7), MAE (1.11 MgCha^{-1}) and coverage (94.7%) thereby making it the best performing model on the premise of the aforementioned validation criteria. The model affords predictive ability little short of the benchmark nominal 95% coverage level (Guhaniyogi and Banerjee, 2019; Sahu, 2022) . We define coverage in this study as defined and interpreted in other studies (Guhaniyogi and Banerjee, 2019) as the percentage of reliably predicted C stock in the study domain. On the other hand, the Landsat-8 based C stock predictive model has 85% (Table 3) coverage for the 95% prediction intervals and higher RMSE and MAE, making it less favourable as the best performing model within the remote sensing data driven Bayesian hierarchical framework.

Fig. 9 (a) and 9 (b) illustrates the scatterplots of observed C stock against the predicted C stock alongside the 95% intervals for both Landsat-8 and Sentinel-2C stock based predictive models. Evidence of the Sentinel-2 based C stock predictive model performing better than its Landsat-8C stock-based counterpart is clear from the scatter plot of the model in Fig. 9 (b). This fact is further bolstered by the Relative Root Mean Square Error (RRMSE) and other validation statistics including the MAE for Sentinel-2 based C stock prediction models illustrated in Table 2.

3.5. MCMC convergence diagnostics

Independent variable coefficients displayed as trace plots for the Landsat-8 and Sentinel-2 based C stock predictions are illustrated in Fig. 10 and Fig. 11. We utilized the trace plots for the aforementioned models as visual assessment of chain convergence. It is clear from the plots that the length N_0, of the burn-in period was sufficiently large to allow Markov chains to converge to the stationary distribution at the end

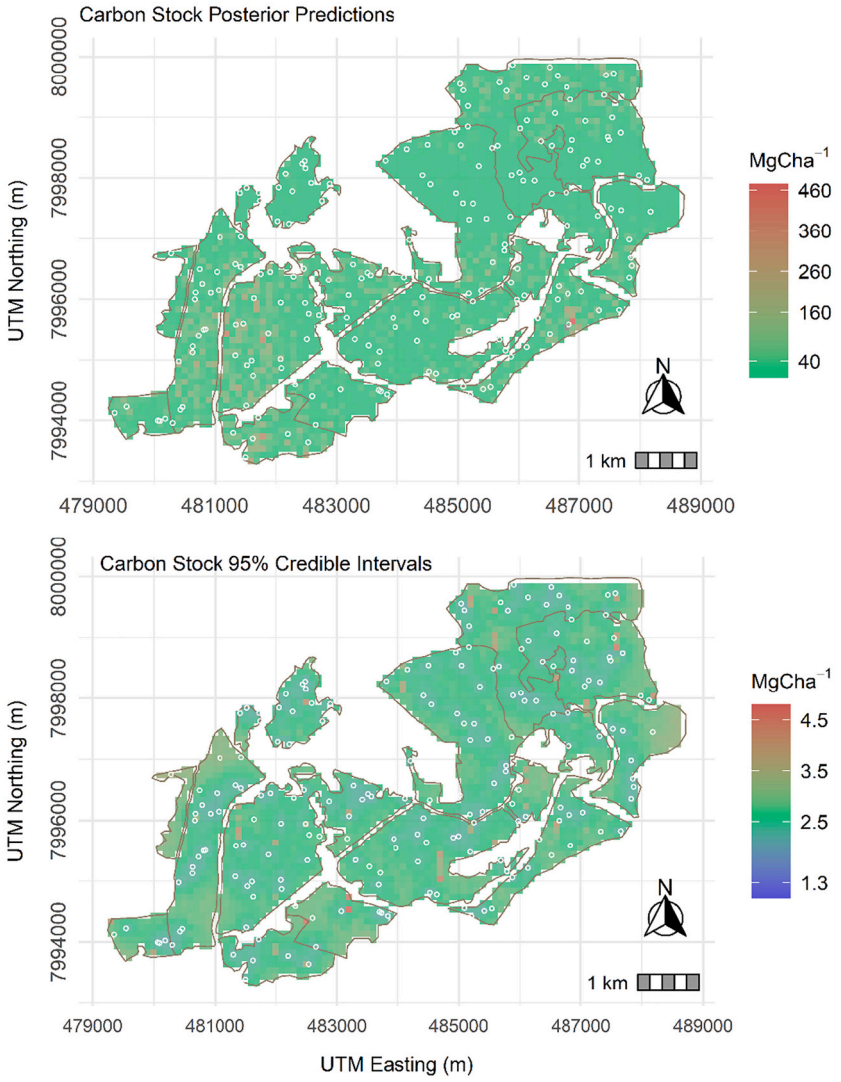


Fig. 7. Landsat-8 based C stock posterior mean and posterior standard deviation.

of the burn-in period (Finley et al., 2007; Gelfand et al., 2004).

For the sake of presentation and improvement of plot aesthetics, we thinned the chains so that only every 10th iteration is retained. Thinning speeds up calculations through the reduction of the Monte Carlo sample size and can also enhance the appearance and presentation of the plots. In both models, the chains seem to have converged to the stationary distribution at the end of the burn-in and hence, to mix fairly well over 5000 iterations (2000 after thinning). It can be deduced from the respective presentations (Fig. 9 and Fig. 10) that the length, N_0 of the MCMC chain should be sufficiently large so that moments and quantiles calculated from the MCMC samples are precise estimates of the resulting characteristics of the posterior (Beloconi and Vounatsou, 2020). A visual assessment of Fig. 9 and Fig. 10 point to a converged sample with MCMC sufficiently mixing well.

4. Discussion

4.1. Landsat-8 and Sentinel-2 derived C stock predictors

The current study juxtaposed findings from Landsat-8 and Sentinel-2 medium resoultion passive satellite sensors in the predictive modelling of C stock using a Bayesian hierarchical approach. We beased the modelling and prediction framework on a set of vegetation indices together with antropogeninc factors as predictors of C stock in a disturbed plantation ecosystem in Zimbabwe. Bordoloi et al. (2022) and Somvanshi and Kumari (2020) have established *NDVI*, *SAVI* and *EVI* to be significant predictors of above ground biomass when using Landsat-8 and Sentinel-2. However, no Bayesian hierarchical approach has yet employed similar predictors from the two medium resolution sensors and established their influence on C stock prediction estimates in a forest plantation set-up.

Apart from the utilization of medium resolution derived covariates, the significance of the anthropogenic predictor variable in the form of distance to nearest settlements in this study is significant. As illustrated in Table 1, distance to nearest settlements and *NDVI* is a significant predictor of C stock in disturbed plantation forest in models estimated from both medium resolution remote sensing sensors. This fact draws logic and support from two main fronts. The plausibility of distance to nearest settlements as a predictor of C stock rationalises from the circumstances prevailing in the plantation forests after the agrarian land reform of 2000. Secondly, studies undertaken elsewhere, including Chinembiri et al. (2013) and Nwobi and Williams (2021), have shown distance to settlements as a significant predictor of C stock in perturbed and related environments.

Previous studies including Finley et al. (2008) proffer interesting exposition on forest biomass over larger scales. The authors mapped AGB using a suite of spatial regression models with Landsat as the source of auxiliary information where residual structure is accounted for by a

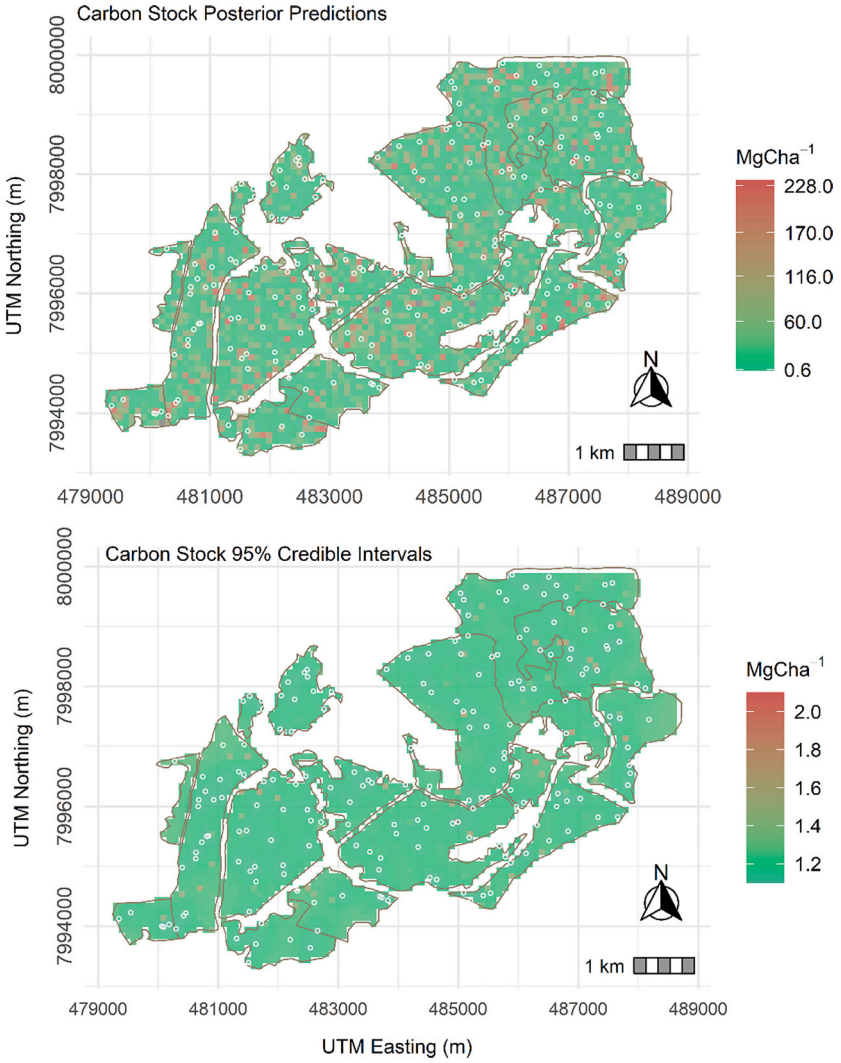


Fig. 8. Sentinel-2 based C stock posterior mean and posterior standard deviation.

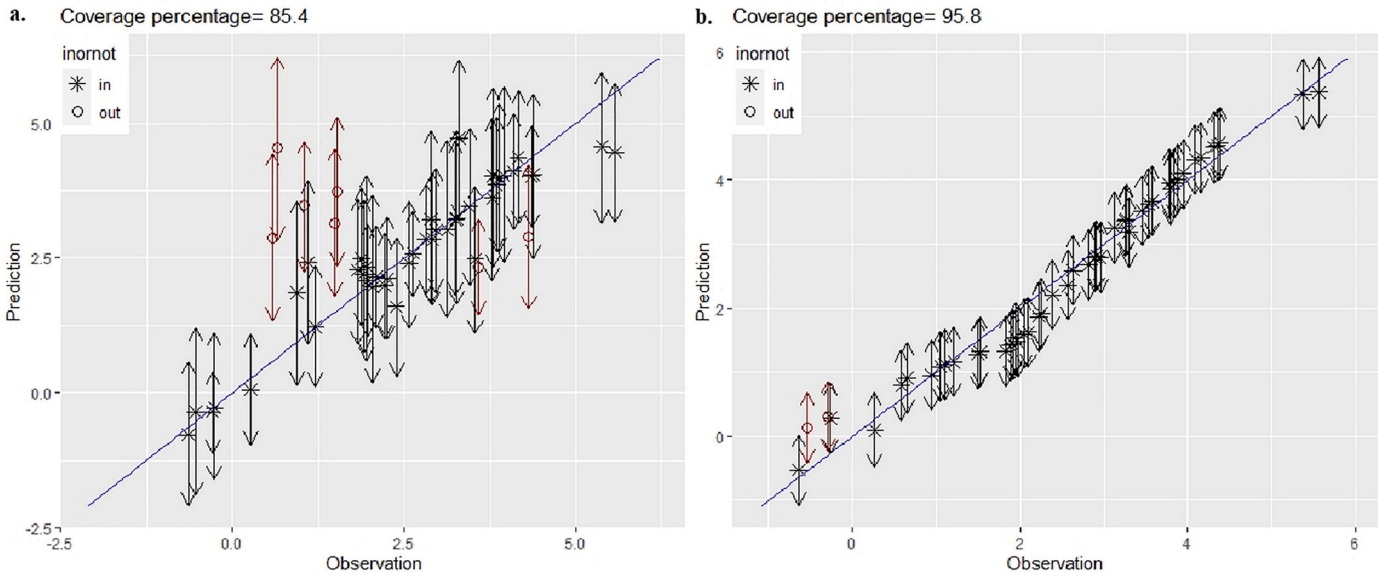


Fig. 9. (a). Spatial model of the Landsat-8C stock-based predictions against observed C stock. (b). Spatial model of the Sentinel-2C stock-based prediction against observed C stock alongside 95% intervals.

Table 2
Landsat-8 and Sentinel-2 derived predictors of C stock.

Parameter	Lansat-8 OLI C stock Model				Sentinel-2 MSI C stock Model			
	Mean	s.d	2.5%	97.5%	Mean	s.d	2.5%	97.5%
Intercept	−2.75	1.01	−4.73	−0.82	−2.41	0.32	−3.04	−1.79
NDVI	2.67	0.99	0.77	4.60	4.97	0.24	4.53	5.45
SAVI	−0.67	0.69	−2.05	0.67	−0.52	0.36	−1.22	0.18
EVI	−0.52	0.53	−1.56	0.53	−0.002	0.097	−0.18	0.20
DIST*	1.47	0.33	0.79	2.09	0.78	0.13	0.52	1.05
σ^2_w	1.24	0.28	0.71	1.78	0.072	0.014	0.045	0.091
σ^2_ε	0.31	0.12	0.079	0.55	0.0047	0.0030	0.0006	0.012
ϕ	0.0017	0.0002	0.0013	0.0019	0.0014	0.0001	0.0013	0.0015

Table 3
Validation statistics for C stock Bayesian hierarchical models.

Model Selection criterion	Landsat 8 Derived Predictors			Sentinel-2 Derived Predictors		
	Independent Error Model	Spatial Intercept only Model	Spatial Model	Independent Error Model	Spatial Intercept only Model	Spatial Model
RMSE ($MgCha^{-1}$)	3.35	2.64	2.69	1.32	3.25	1.16
RRMSE	0.11	0.09	0.09	0.04	0.11	0.04
MAE ($MgCha^{-1}$)	2.44	1.70	1.79	1.25	2.16	1.11
CRPS ($MgCha^{-1}$)	2.18	1.46	1.48	1.85	1.75	1.13
CVG (%)	91.7	85.4	85.4	95.8	89.6	94.7
DIC	210.2	56.9	43.1	−244.4	37.8	−554.7

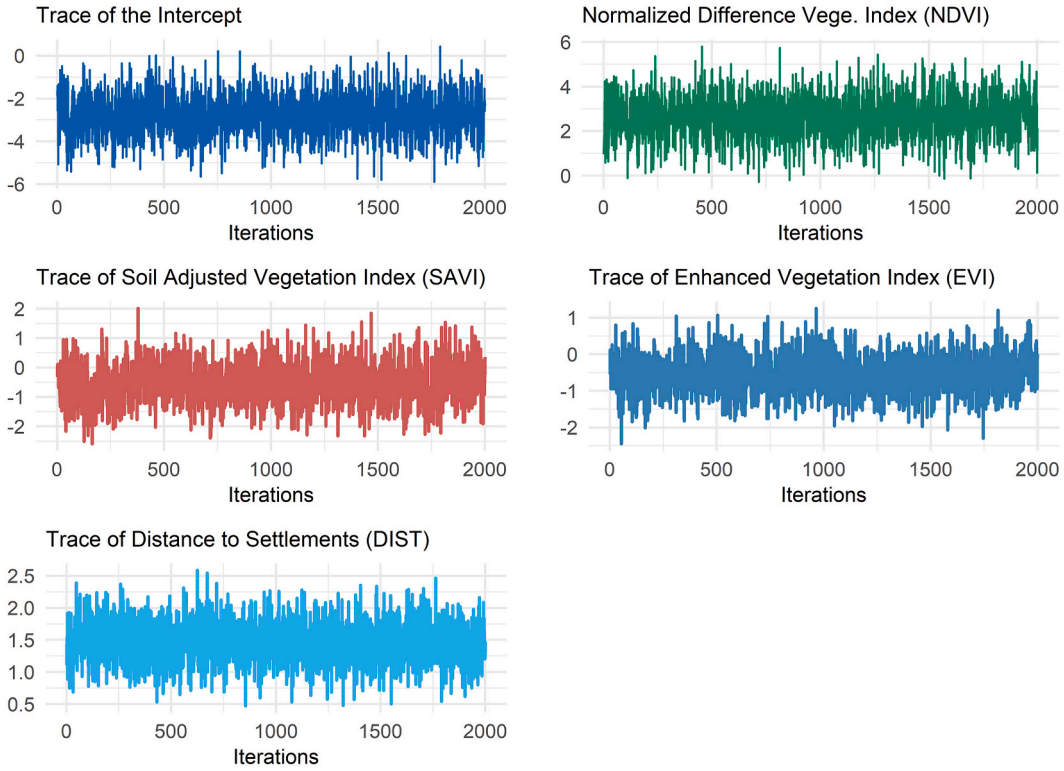


Fig. 10. Lansat-8 based C stock prediction MCMC trace plots.

Gaussian process that has the capacity to account for short and long range spatially structured dependence. Their findings demonstrated effective long-range dependence to the tune of 8000 m with associated 95% CIs of 4000 to 11,000 m. Similar improvements in spatially structured dependence is also reported in Datta et al. (2016) work when modelling AGB in the conterminous United States.

4.2. Parameter uncertainty and model performance of new generation sensor-based C stock models

An investigation of the AGB distribution of a mangrove forest in Vietnam employing a combination of remote sensing data and Artificial Neural Network (ANN) gave mangrove AGB predictions that ranged from 6.53 to 368.2 $Mgha^{-1}$ in 2000 and from 13.75 to 320.3 $Mgha^{-1}$ in 2020, respectively (Do et al., 2022). AGB predictions made by Fararoda et al. (2021) using machine learning methods including AdaBoost,

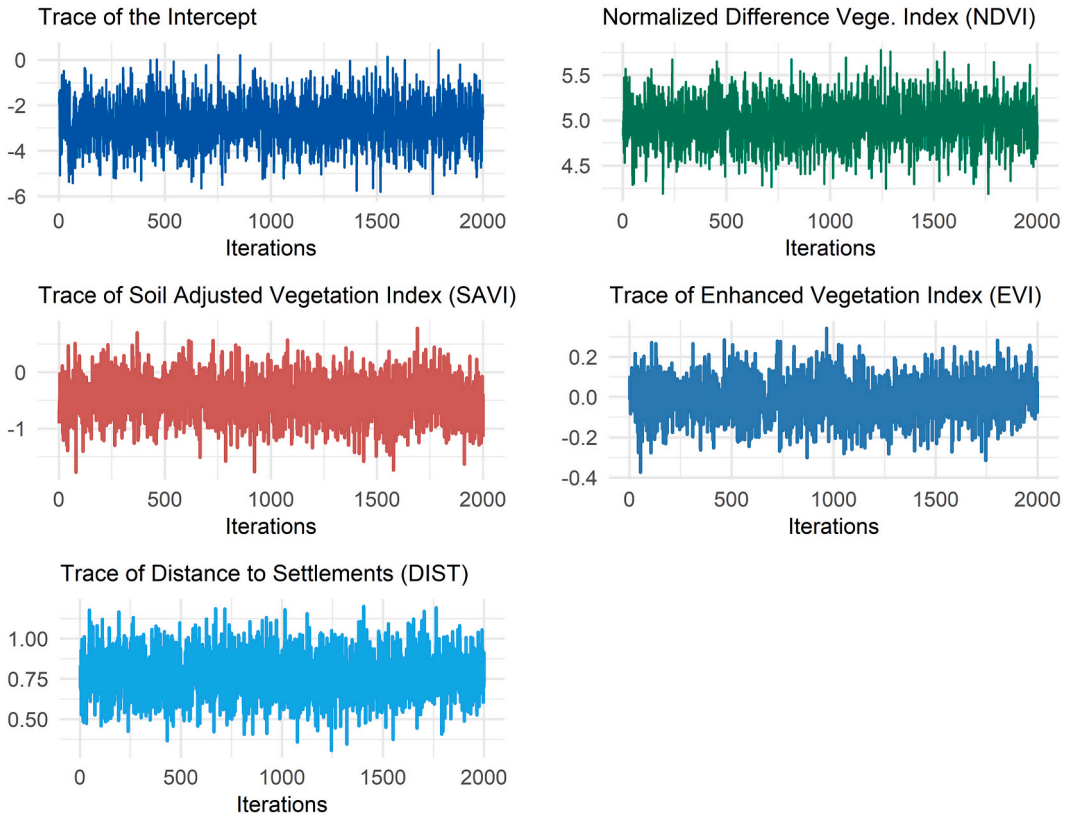


Fig. 11. Sentinel-2 based C stock prediction MCMC trace plots.

random decision forest, multilayer neural networks and Bayesian ridge regression endorsed random forest and AdaBoost as the best performing methods. The type of sensor and the method of prediction used have great influence over AGB model performance. For instance, Fassnacht et al. (2014) compared LiDAR and hyperspectral data for modelling AGB and concluded that LiDAR data fused with Random forest models offered the best AGB model performance.

In terms of the accuracy of prediction, it is vital to compare the results of the present study with recent publications in literature from similar and related terrestrial biomes. Jiang et al. (2021) established an average AGB RMSE of 40.92 $MgCha^{-1}$ in northeast China whilst Dang et al. (2019) reported a RMSE of 36.67 $MgCha^{-1}$ in Vietnam using Random Forest (RF). Takagi et al. (2015) utilized LiDAR for predicting forest biomass in Hokkaido, Japan, and established a biomass RMSE prediction of 19.10 $MgCha^{-1}$. Differences in the prediction accuracy between the reported results in literature and our study can be partly explained by the differences in forest density and largely by the differences in the modelling approaches between results reported in literature and our methodology. First, the aforementioned studies are likely to have significant and higher AGB density than in Zimbabwe as they were undertaken in tropical and subtropical rainforest biomes. Secondly and most importantly, the present study uses a hierarchical Bayesian approach which has the advantage of having access to the entire posterior predictive distribution (Goulard and Voltz, 1992).

Location based C stock uncertainty maps presented in this study are needed for monitoring ecosystem health and areas of interest in the plantation forests for interventions during times of extreme weather, and in times of disease outbreaks. Such information provides timely decision making and management by forest practitioners. With accuracies for both Landsat-8 and Sentinel-2 based models exceeding accuracies established in other similar environments (Ahmed et al., 2022; Fan et al., 2022; Jiang et al., 2021; Takagi et al., 2015; Xiong and Wang, 2022), the application of results of the present study can therefore be utilized for comprehensive and location-based ecological interventions. The Sentinel-2 based C stock predictive model yielded the shortest CIWs, implying more precise C stock predictions than the Landsat-8 based C stock estimates. Hence, the usefulness of Landsat-8 based C stock predictive model is less attractive as it cannot match the standards and qualities inherent in the improved Sentinel-2 for similar modelling framework as the current study. On the other hand, the Landsat-8 based C stock predictive model gave slightly higher RMSE (2.69 $MgCha^{-1}$) and slightly lower coverage (85.4%) than its Sentinel-2 based C stock spatial predictive counterpart. This also means monitoring and conservation for extensive forest ecosystems can be done with greater accuracy and precision, thereby improving forest carbon accounting, especially in inaccessible areas where samples cannot be easily taken. Higher prediction accuracy means mapping and accounting for natural resources can be achieved with great accuracy and save forest species from extinction in the event of disease outbreaks and extreme environmental conditions (Ahmed et al., 2022; Do et al., 2022; Fan et al., 2022).

As attributed in Frampton et al., 2013 and Gerald et al., 2017, the posterior predictive distribution of Sentinel-2 based C stock predictions are more precise than their Landsat-8 based counterparts as they have far shorter 95% CIWs. This implies that highly certain predictions of C stock in disturbed plantation ecosystems can better be made with Sentinel-2 derived vegetation indices (*NDVI*) and distance to nearest settlements than with similar spectral and anthropogenic variables derived from the Landsat-8 based C stock hierarchical model. The quality of predictors deriving from the tested medium resolution sensors is a critical point of discussion in this case (Ahmed and De Marsily, 1987). More refined spectral and spatial resolution (10 m, 20 m) of Sentinel-2 derived spectral indices offer an opportunity for enhanced and accurate monitoring of forest resources (Mutanga et al., 2016; Somvanshi and Kumari, 2020). Our results fall in the same realm as those of Korhonen et al. (2017), Astola et al. (2019) and Jha et al. (2021) due to the differences in spatial and spectral properties of the two

sensors. As such, despite the departure in the modelling methodology used by the current study, comparison of the predictive performance of Sentinel-2 based C stock model still outperforms its Landsat-8 counterpart and other machine learning methods reported in literature. This can be attributed to the Bayesian hierarchical modelling approach utilized in the present study, which has the strength of incorporating multiple sources of uncertainty into the prediction modelling framework (Azevedo, 2021). Furthermore, Dang et al. (2019) utilized Sentinel-2 together with field based biomass data using random forest (RF) and machine learning algorithm for estimating AGB in Vietnam. Results employing the RF algorithm together with Sentinel-2 derived auxiliary information resulted in more accurate predictions of AGB compared to the machine learning methods. Consequently, findings in literature using Sentinel-2 as a data source have shown the former to be superior than other data sources, thereby making Sentinel-2 medium resolution imagery a data source of choice for forests management and ecological assessment.

Wu et al. (2016) and Xiong and Wang (2022) compared machine learning methods for AGB estimation based on Landsat imagery and established Random forest to be the best performing method with a RMSE of 26.44 tons/ha over the other methods like support vector regression, k-nearest neighbour and stochastic gradient boosting. As noted in the present study, hierarchical Bayesian geostatistical methods offer superior C stock predictions than the other methods in literature. Nevertheless, hierarchical geostatistical methods cannot be easily adopted and utilized by forest practitioners for ecological monitoring. Their complexity and the lack of readily available easy to use application packages greatly limit their adoption and application in ecological monitoring. The fusion of medium and high resolution satellite sensors greatly enhance the accuracy of predicting forest biomass (Fararoda et al., 2021). However, differences in model performance reported in literature and the present study are particularly significant as they suggest the selection of appropriate statistical and modelling methods to be more effective than putting more effort and resources in field data collection. Previous research comparing Sentinel-2 and Landsat-8 sensors noted improvements in the spatial and spectral capabilities of Sentinel-2 in discriminating rangeland management practices (Sibanda et al., 2017), estimating Leaf Area Index (LAI) (Korhonen et al., 2017) and to enhance the quality of classification of built-up areas (Pesaresi et al., 2016). The inclusion of three bands within the red-edge portion of the electromagnetic spectrum is given as the rationale for Sentinel-2's superior performance in all the comparative studies. Hence, the performance of the Sentinel-2 based C stock predictive model in this research can be traced to the additional and narrower band channels, in particular, the red-edge band coupled with improved spatial resolution. However, studies that have utilized Landsat-8 in isolation to Sentinel-2 for predicting AGB, have noted significant improvement in the spectral capabilities of the imagery.

Predictions in the studied region of Zimbabwe show a much reduced density of C stock compared with other tropical and sub-tropical forest ecosystems in similar environments like the Indian forests, with average biomass density of more than 420 $Mgha^{-1}$ (Fararoda et al., 2021). Implementation of conservation practices needed for ecological restoration of the disturbed plantation forest ecosystems therefore becomes an urgent matter. Depressed C stock concentration in these environments may also imply that the plantation forest may take much longer to return to their former state following perturbations from farming and gold panning activities of settlers in the plantations.

5. Conclusions

The goal of our investigation was to develop and test the performance of a Bayesian hierarchical modelling approach using C stock as the outcome variable, under the influence of different but related medium resolution satellite sensor data sources. By comparison of models with varying levels of richness from Landsat-8 and Sentinel-2, we were able to demonstrate that the Sentinel-2 based C stock spatial hierarchical model has the best predictive characteristics in terms of the *RMSE*, *MAE*, *CRPS* and *DIC*, and is more attractive as compared to the Landsat-8 based C stock hierarchical model. For the sake of our goal, the Sentinel-2 based C stock spatial predictive model is the best as it is applicable to domains beyond the training sample. The Sentinel-2 based C stock spatial predictive model is more useful than the Landsat-8 based C stock based predictive model in assisting plantation forest practitioners. This is critical as forest practitioners should correctly advise national governments on the potential of resuscitated plantation forests in sequestering carbon in the country. The findings of this study will aid in the understanding of dynamics between vegetation indices and C stock concentration and determining area specific sites for intervention for ecological restoration and monitoring. The high cost of obtaining data for ecological monitoring and validation of biomass models by forest practitioners and managers can be substantially reduced as readily available and high-resolution Sentinel-2 remote sensing data is cheap to acquire.

There is however, a need to extend the present modelling framework to a broader coverage of auxiliary information by integrating broadband vegetation indices and bioclimatic variables in a comprehensive Bayesian hierarchical framework. The high residual spatial dependence demonstrated in Sentinel-2 based C stock hierarchical modelling implies the effectiveness of covariate information from remote sensing. Over and above the significant reduction in C stock prediction uncertainty when modelling with covariates from finer spectral and spatial resolution, our model offers the possibility of robust modelling if we adopt a framework encompassing a much broader suite of independent variables to offer better predictive maps of C stock distribution in disturbed plantation ecosystems. We conclude that Sentinel-2 data can be proposed as the primary earth observation data source in the management and monitoring of managed ecosystems with history of ecological fragility.

Declaration of Competing Interest

None.

Data availability

Data will be made available on request.

Acknowledgments

The authors express their gratitude to Mr. Kutsaranga, Dr. Matowanyika and Mr. Mukwekwe for providing access to plantations of the sampled study area at Lot 75A of Nyanga Downs in Manicaland Province. This research was supported by the University of KwaZuluNatal, college of Agricultural, Earth & Environmental Science.

References

Agarwal, D.K., et al., 2005. Tropical deforestation in Madagascar: analysis using hierarchical, spatially explicit, Bayesian regression models. Ecol. Model. 185 (1), 105–131. https://doi.org/10.1016/j.ecolmodel.2004.11.023.

Ahmed, De Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resour. Res. 23 (9), 1717–1737.

Ahmed, N., Atzberger, C., Zewdie, W., 2022. The potential of modeling Prosopis Juliflora invasion using Sentinel-2 satellite data and environmental variables in the dryland ecosystem of Ethiopia. Ecol. Informat. 68, 101545 https://doi.org/10.1016/j.ecoinf.2021.101545.

Astola, H., et al., 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. Remote Sens. Environ. 223, 257–273. https://doi.org/10.1016/j.rse.2019.01.019.

Azevedo, L., 2021. Model reduction in geostatistical seismic inversion with functional data analysis. Geophysics 87 (1), M1–M11. https://doi.org/10.1190/geo2021-0096.1. Society of Exploration Geophysicists.

Babcock, C., et al., 2012. Multivariate spatial regression models for predicting individual tree structure variables using LiDAR data. IEEE J. Select. Top. Appl. Earth Observat. Rem. Sens. 6 (1), 6–14. IEEE.
Babcock, C., et al., 2016. Modeling forest biomass and growth: coupling long-term inventory and LiDAR data. Remote Sens. Environ. 182, 1–12. https://doi.org/10.1016/j.rse.2016.04.014.
Babcock, C., et al., 2018. Geostatistical estimation of forest biomass in interior Alaska combining Landsat-derived tree cover, sampled airborne lidar and field observations. Remote Sens. Environ. 212, 212–230. https://doi.org/10.1016/j.rse.2018.04.044.
Babcock, C., Finley, A.O., Bradford, J.B., Kolka, R., Birdsey, R., Ryan, M.G., 2015. LiDAR based prediction of forest biomass using hierarchical models with spatially varying coefficients. Remote. Sens. Environ. 169, 113–127.
Baloloy, A.B., et al., 2018. Estimation of mangrove forest aboveground biomass using multispectral bands, vegetation indices and biophysical variables derived from optical satellite imageries: rapideye, planetscope and sentinel-2. In: ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. Copernicus Publications, IV–3, pp. 29–36. https://doi.org/10.5194/isprs-annals-IV-3-29-2018.
Baugh, W.M., Groeneveld, D.P., 2006. Broadband vegetation index performance evaluated for a low-cover environment. Int. J. Remote Sens. 27 (21), 4715–4730.
Beloconi, A., Vounatsou, P., 2020. Bayesian geostatistical modelling of high-resolution NO2 exposure in Europe combining data from monitors, satellites and chemical transport models. Environ. Int. 138, 105578 https://doi.org/10.1016/j.envint.2020.105578.
Bordoloi, R., et al., 2022. Satellite based integrated approaches to modelling spatial carbon stock and carbon sequestration potential of different land uses of Northeast India. Environm. Sustainabil. Indicat. 13, 100166 https://doi.org/10.1016/j.indic.2021.100166.
Box, G.E.P., Cox, D.R., 1964. An analysis of transformations (with discussion). J. R. Stat. Soc. 26, 211–252.
Box, G.E.P., Cox, D.R., 1982. An analysis of transformations revisited, rebutted, 77, 209–210. J. Am. Stat. Assoc. 77, 209–210.
Brown, S., 1997. Estimating biomass and biomass change of tropical forests: a primer. In: FAO Forestry Paper, 134.
Brus, D.J., De Gruijter, J.J., Van Groenigen, J.W., 2006. Chapter 14 designing spatial coverage samples using the k-means clustering algorithm. Dev. Soil Sci. 31, 183–192.
Chinembiri, T.S., et al., 2013. The precision of C stock estimation in the Ludhikola watershed using model-based and design-based approaches. Nat. Resour. Res. 22 (4), 297–309. https://doi.org/10.1007/s11053-013-9216-6.
Clerici, N., et al., 2016. Estimating aboveground biomass and carbon stocks in Periurban Andean secondary forests using very high resolution imagery. Forests. https://doi.org/10.3390/f7070138.
Cochran, W.G., 1977. Sampling Techniques. John Wiley & Sons, New York.
Cressie, N., 1993. Statistics for Spatial Data. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons.
Cross, W.F., Rosi-Marshall, E.J., Behn, K.E., Kennedy, T.A., Hall, R.O., Fuller, A.E., Baxter, C.V., 2010. Invasion and production of New Zealand mud snails in the Colorado River, Glen Canyon. Biol. Invasions 12 (9), 3033–3043.
Dang, A.T.N., et al., 2019. Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park Vietnam. Ecol. Informat. 50, 24–32. https://doi.org/10.1016/j.ecoinf.2018.12.010.
Datta, A., et al., 2016. Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. J. Am. Stat. Assoc. 111 (514), 800–812. https://doi.org/10.1080/01621459.2015.1044091. Taylor & Francis.
Demirhan, H., Kalaylioglu, Z., 2015. Joint prior distributions for variance parameters in Bayesian analysis of normal hierarchical models. J. Multivar. Anal. 135, 163–174. https://doi.org/10.1016/j.jmva.2014.12.013.
Diggle, P., Ribeiro Jr., P., 2007. Model-based geostatistics. J. R. Stat. Soc. C. https://doi.org/10.1007/978-0-387-48536-2.
Do, A.N.T., et al., 2022. Monitoring landscape fragmentation and aboveground biomass estimation in can Gio mangrove biosphere reserve over the past 20 years. Ecol. Informat. 70, 101743 https://doi.org/10.1016/j.ecoinf.2022.101743.
Drusch, M., et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.
Duchene, S., et al., 2016. Cross-validation to select Bayesian hierarchical models in phylogenetics. BMC Evol. Biol. 16 https://doi.org/10.1186/s12862-016-0688-y.
Fan, M., Wang, X., Yang, G., 2022. Spatial characteristics of vegetation habitat suitability and mountainous settlements and their quantitative relationships in upstream of Min River, southwestern of China. Ecol. Informat. 68, 101541 https://doi.org/10.1016/j.ecoinf.2021.101541.
Fararoda, R., et al., 2021. Improving forest above ground biomass estimates over Indian forests using multi source data sets with machine learning algorithm. Ecol. Informat. 65, 101392 https://doi.org/10.1016/j.ecoinf.2021.101392.
Fassnacht, F.E., et al., 2014. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. Remote Sens. Environ. 154, 102–114. https://doi.org/10.1016/j.rse.2014.07.028.
Finley, A.O, Sang, H., Banerjee, S., Gelfand, A.E., 2009. Improving the performance of predictive process modeling for large datasets. Comput. Stat. Data Anal. 53 (8), 2873–2884. https://doi.org/10.1016/j.csda.2008.09.008. PMID: 20016667; PMCID: PMC274316.
Finley, A., Sudipto, B., Carlin, B., 2007. spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. J. Stat. Softw. 19 https://doi.org/10.18637/jss.v019.i04.
Finley, A.O., Banerjee, S., McRoberts, R.E., 2008. A Bayesian approach to multi-source forest area estimation. Environ. Ecol. Stat. 15, 241–258.
Finley, A.O., Banerjee, S., MacFarlane, D.W., 2011. A hierarchical model for quantifying forest variables over large heterogeneous landscapes with uncertain Forest areas. J. Am. Stat. Assoc. 106 (493), 31–48. https://doi.org/10.1198/jasa.2011.ap09653. Taylor & Francis.
Frampton, W.J., et al., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS J. Photogramm. Remote Sens. 82, 83–92. https://doi.org/10.1016/j.isprsjprs.2013.04.007.
FAO, 2018. https://agris.fao.org/agris-search/search.do;jsessionid=5426E3747CA068F916F1DB81467F9426?request_locale=fr&recordID=US201800026662&query=&sourceQuery=&sortField=&sortOrder=&agrovocString=&advQuery=¢erString=&enableField=.
Gelfand, A.E., 2012. Hierarchical modeling for spatial data problems. Spat. Stat. 1, 30–39. https://doi.org/10.1016/j.spasta.2012.02.005.
Gelfand, A.E., et al., 2004. Nonstationary multivariate process modeling through spatially varying coregionalization. Test 13 (2), 263–312. https://doi.org/10.1007/BF02595775.
Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal. 1 (3), 515–534. https://doi.org/10.1214/06-BA117A.
Gerald, F., et al., 2017. Lansat-8 vs. sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. GISci. Rem. Sens. 1–26.
Gibbs, H.K., et al., 2007. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. Environ. Res. Lett. 2 (4), 45023. https://doi.org/10.1088/1748-9326/2/4/045023. IOP Publishing.
González-Vélez, J.C., et al., 2021. An artificial intelligent framework for prediction of wildlife vehicle collision hotspots based on geographic information systems and multispectral imagery. Ecol. Informat. 63, 101291 https://doi.org/10.1016/j.ecoinf.2021.101291.
Goulard, M., Voltz, M., 1992. Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. Math. Geol. 24 (3), 269–286. https://doi.org/10.1007/BF00893750.
Green, E.J., Finley, A.O., Strawderman, W.E., 2020. Introduction to Bayesian Methods in Ecology and Natural Resources. Springer, Cham. https://doi.org/10.1007/978-3-030-60750-0.
Gregoire, T.G., 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Ca. J. For. Res 28, 1429–1447.
Guhaniyogi, R., Banerjee, S., 2019. Multivariate spatial meta kriging. Stat. Probabil. Lett. 144, 3–8. https://doi.org/10.1016/j.spl.2018.04.017.
Hengl, T., et al., 2004. A double continuous approach to visualization and analysis of categorical maps. Int. J. Geogr. Inf. Sci. 18 (2), 183–202. https://doi.org/10.1080/13658810310001620924. Taylor & Francis.
IPCC, 2006. Good Practice Guidance for Land Use, Land-Use Change and Forestry. Institute for Global Environmental Strategies, Kanagawa, Japan.
Jackman, S., 2000. Estimation and inference via Bayesian simulation: an introduction to Markov Chain Monte Carlo. Am. J. Polit. Sci. 44 (2), 375–404. https://doi.org/10.2307/2669318 [Midwest Political Science Association, Wiley].
Jha, N., et al., 2021. The real potential of current passive satellite data to map aboveground biomass in tropical forests. Rem. Sens. Ecol. Conservat. 7 https://doi.org/10.1002/rse2.203.
Jiang, F., et al., 2021. Estimating the aboveground biomass of coniferous forest in Northeast China using spectral variables, land surface temperature and soil moisture. Sci. Total Environ. 785, 147335 https://doi.org/10.1016/j.scitotenv.2021.147335.
Johnson, K.D., et al., 2014. Integrating forest inventory and analysis data into a LIDAR-based carbon monitoring system. Carbon Balance Manag. 9 (1), 3. https://doi.org/10.1186/1750-0680-9-3.
Korhonen, L., et al., 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. Remote Sens. Environ. 195, 259–274. https://doi.org/10.1016/j.rse.2017.03.021.
Leenhouts, B., 1998. Assessment of biomass burning in the conterminous United States. Conserv. Ecol. 2 (1).
Lefsky, M.A., 2010. A global forest canopy height map from the moderate resolution imaging Spectroradiometer and the geoscience laser altimeter system. Geophys. Res. Lett. 37 (15). Wiley Online Library.
Li, C., Li, X., 2019. Hazard rate and reversed hazard rate orders on extremes of heterogeneous and dependent random variables. Stat. Probabil. Lett. 146, 104–111. https://doi.org/10.1016/j.spl.2018.11.005.
Millington, A., Townsend, J., 1989. Biomass Assessment, 1st ed. Routledge. https://doi.org/10.4324/9781315066417.
Mutanga, O., Dube, T., Ahmed, F., 2016. Progress in remote sensing: vegetation monitoring in South Africa. S. Afr. Geogr. J. 98 (3), 461–471. https://doi.org/10.1080/03736245.2016.1208586. Routledge.
Newsday, 2017. Forestry Commision decentralise issuance of timber movement. Retrieved April 25, 2019, from. https://www.newsday.co.zw/2017/07/forestry-commision-decentralise-issuance-timber-movement/.
Nwobi, C.J., Williams, M., 2021. Natural and anthropogenic variation of stand structure and aboveground biomass in Niger Delta Mangrove Forests. Front. For. Global Change 4, 1–15. Available at. https://www.frontiersin.org/article/10.3389/ffgc.2021.746671.
Pascual, A., Tupinambá-Simões, F., de Conto, T., 2022. Using multi-temporal tree inventory data in eucalypt forestry to benchmark global high-resolution canopy height models. A showcase in Mato Grosso, Brazil. Ecol. Informat. 70, 101748 https://doi.org/10.1016/j.ecoinf.2022.101748.

Pesaresi, M., et al., 2016. Assessment of the added-value of Sentinel-2 for detecting built-up areas. Remote Sens. https://doi.org/10.3390/rs8040299.
R Core Development, T, 2008. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Ranghetti, L., et al., 2020. "sen2r": an R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data. Comput. Geosci. 139, 104473 https://doi.org/10.1016/j.cageo.2020.104473.
Ravindranath, N.H., Ostwald, M., 2008. Carbon inventory methods handbook for greenhouse gas inventory, carbon mitigation and Roundwood production projects. Adv. Global Change Res. 29 https://doi.org/10.1007/978-1-4020-6547-7_1.
Sahu, S.K., 2022. Bayesian Modeling of Spatio Temporal Data with R, 1st ed. Chapman and Hall/CRC, Highfield, Southampton, UK. https://doi.org/10.1201/9780429318443.
Semela, M., Ramoelo, A., Adelabu, S., 2020. Testing and comparing the applicability of sentinel-2 and landsat 8 reflectance data in estimating mountainous herbaceous biomass before and after fire using random forest modelling. In: IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, pp. 4493–4496. https://doi.org/10.1109/IGARSS39084.2020.9323446.
Shumba, E.M., Marongwe, D., 2016. The convention on biological diversity: an overview and lessons learnt from the Zimbabwean experience. Int. For. Rev. 31–77.
Sibanda, M., et al., 2017. Estimating biomass of native grass grown under complex management treatments using WorldView-3 spectral derivatives. Remote Sens. https://doi.org/10.3390/rs9010055.
Simard, M., et al., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. Biogeosci. 116 (G4). Wiley Online Library.
Somvanshi, S.S., Kumari, M., 2020. Comparative analysis of different vegetation indices with respect to atmospheric particulate pollution using sentinel data. Appl. Comput. Geosci. 7, 100032 https://doi.org/10.1016/j.acags.2020.100032.
Song, C., et al., 2010. Estimating average tree crown size using spatial information from Ikonos and QuickBird images: across-sensor and across-site comparisons. Remote Sens. Environ. 114, 1099–1107. https://doi.org/10.1016/j.rse.2009.12.022.
Sovdat, B., Kadunc, M., Batič, M., Milčinski, G., 2019. Natural color representation of Sentinel-2 data. Remote Sens. Environ. 225, 392–402. https://doi.org/10.1016/j.rse.2019.01.036.
Stoyan, Y., Yaskov, G., 2014. Packing unequal circles into a strip of minimal length with a jump algorithm. Optimization Letters 8 (3), 949–970.
Takagi, K., et al., 2015. Forest biomass and volume estimation using airborne LiDAR in a cool-temperate forest of northern Hokkaido Japan. Ecol. Informat. 26, 54–60. https://doi.org/10.1016/j.ecoinf.2015.01.005.
Thompson, S.K., 2002. On sampling and experiments. Environmetrics 13 (5–6), 429–436. https://doi.org/10.1002/env.532.
Traore, M., Tieguhong, J., 2018. How Forestry Contributes to the African Development Bank's 'High 5s' Priorities for Africa: Challenges and Opportunities.
Ver Hoff, J., 2002. Sampling and Geostatistics for Spatial Data. Écoscience 9 (2), 152–161. http://www.jstor.org/stable/42901479.
Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Comput. Geosci. 36 (10), 1261–1267. https://doi.org/10.1016/j.cageo.2010.04.005.
Wang, Q., et al., 2020. Comparative analysis of Landsat-8, Sentinel-2, and GF-1 data for retrieving soil moisture over wheat farmlands. Remote Sens. https://doi.org/10.3390/rs12172708.
Wang, Q., Ni, J., Tenhunen, J., 2005. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. Glob. Ecol. Biogeogr. 14, 379–393. https://doi.org/10.1111/j.1466-822X.2005.00153.x.
Wu, C., et al., 2016. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. J. Appl. Remote. Sens. 10 (3), 35010. https://doi.org/10.1117/1.JRS.10.035010.
Xiong, Y., Wang, H., 2022. Spatial relationships between NDVI and topographic factors at multiple scales in a watershed of the Minjiang River, China. Ecol. Informat. 69, 101617 https://doi.org/10.1016/j.ecoinf.2022.101617.
Zunguze, A.X., 2012. Quantificação de carbono sequestrado em povoamentos de eucalyptus spp na floresta de Inhamacari-Manica. Universidade Eduardo Mondlane.