Multisource forest inventories: A model-based approach using k-NN to reconcile forest attributes statistics and map products

Ankit Sagar, Cédric Vega, Olivier Bouriaud, Christian Piedallu, Jean-Pierre Renaud

ISPRS Journal of Photogrammetry and Remote Sensing 192 (2022) 175–188

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Multisource forest inventories: A model-based approach using *k*-NN to reconcile forest attributes statistics and map products

Ankit Sagar [a,b,c], Cédric Vega [b,*], Olivier Bouriaud [b,d], Christian Piedallu [c], Jean-Pierre Renaud [b,e]

[a] Université de Lorraine, Faculté des Sciences et Technologies - Campus Aiguillettes, 54506 Vandœuvre Les Nancy, France
[b] Laboratoire d'Inventaire Forestier, Université de Lorraine, ENSG, IGN, INRAe, 54000 Nancy, France
[c] UMR SILVA INRAe-AgroParisTech-Université de Lorraine, 54000 Nancy, France
[d] Ștefan cel Mare University of Suceava, 13 University street, Suceava 720229, Romania
[e] Office National des Forêts, Pôle Recherche Développement Innovation, 54600 Villers-lès-Nancy, France

ARTICLE INFO

Keywords:
Multisource forest inventories
Statistical inference
Forest attribute maps
Reliability assessment
LiDAR
Photogrammetry

ABSTRACT

Forest map products are widely used and have taken benefit from progresses in the multisource forest inventory approaches, which are meant to improve the precision of forest inventory estimates at high spatial resolution. However, estimating errors of pixel-wise predictions remains difficult, and reconciling statistical outcomes with map products is still an open and important question.

We address this problem using an original approach relying on a model-based inference framework and *k*-nearest neighbours (*k*-NN) models to produce pixel-wise estimations and related quality assessment. Our approach takes advantage of the resampling properties of a model-based estimator and combines it with geometrical convex-hull models to measure respectively the precision and accuracy of pixel predictions. A measure of pixel reliability was obtained by combining precision and accuracy.

The study was carried out over a 7,694 km^2 area dominated by structurally complex broadleaved forests in centre of France. The targeted forest attributes were growing stock volume, basal area and growing stock volume increment. A total of 819 national forest inventory plots were combined with auxiliary data extracted from a forest map, Landsat 8 images, and 3D point clouds from both airborne laser scanning and digital aerial photogrammetry. *k*-NN models were built independently for both 3D data sources. Both selected models included 5 auxiliary variables, and were generated using 5 neighbours, and most similar neighbours distance measure. The models showed relative root mean square error ranging from 35.7% (basal area, digital aerial photogrammetry) in calibration to 63.4% (growing stock volume increment, airborne laser scanning) in the validation set. At pixel level, we found that a minimum of 86.4% of the predictions were of high precision as their bootstrapped coefficient of variation fall below calibration's relative root mean square error. The amount of extrapolation varied from 4.3% (digital aerial photogrammetry) to 6.3% (airborne laser scanning). A relationship was found between extrapolation and *k*-NN distance, opening new opportunities to correct extrapolation errors. At the population level, airborne laser scanning and digital aerial photogrammetry performed similarly, offering the possibility to use digital aerial photogrammetry for monitoring purposes.

The proposed method provided consistent estimates of forest attributes and maps, and also provided spatially explicit information about pixel predictions in terms of precision, accuracy and reliability. The method therefore produced high resolution outputs, significant for either decision making or forest management purposes.

1. Introduction

National Forest Inventories (NFIs) are crucial forest assessment programmes, providing estimates about the amount and state of the forest resources, and information forming the basis for the development of various forest management policies (Breidenbach et al., 2021). NFIs

* Corresponding author.
E-mail addresses: ankit.sagar@univ-lorraine.fr (A. Sagar), cedric.vega@ign.fr (C. Vega), olivier.bouriaud@ign.fr, obouriaud@usm.ro (O. Bouriaud), christian.piedallu@agroparistech.fr (C. Piedallu), jean-pierre.renaud-02@onf.fr (J.-P. Renaud).

https://doi.org/10.1016/j.isprsjprs.2022.08.016
Received 11 February 2022; Received in revised form 14 August 2022; Accepted 17 August 2022
Available online 26 August 2022

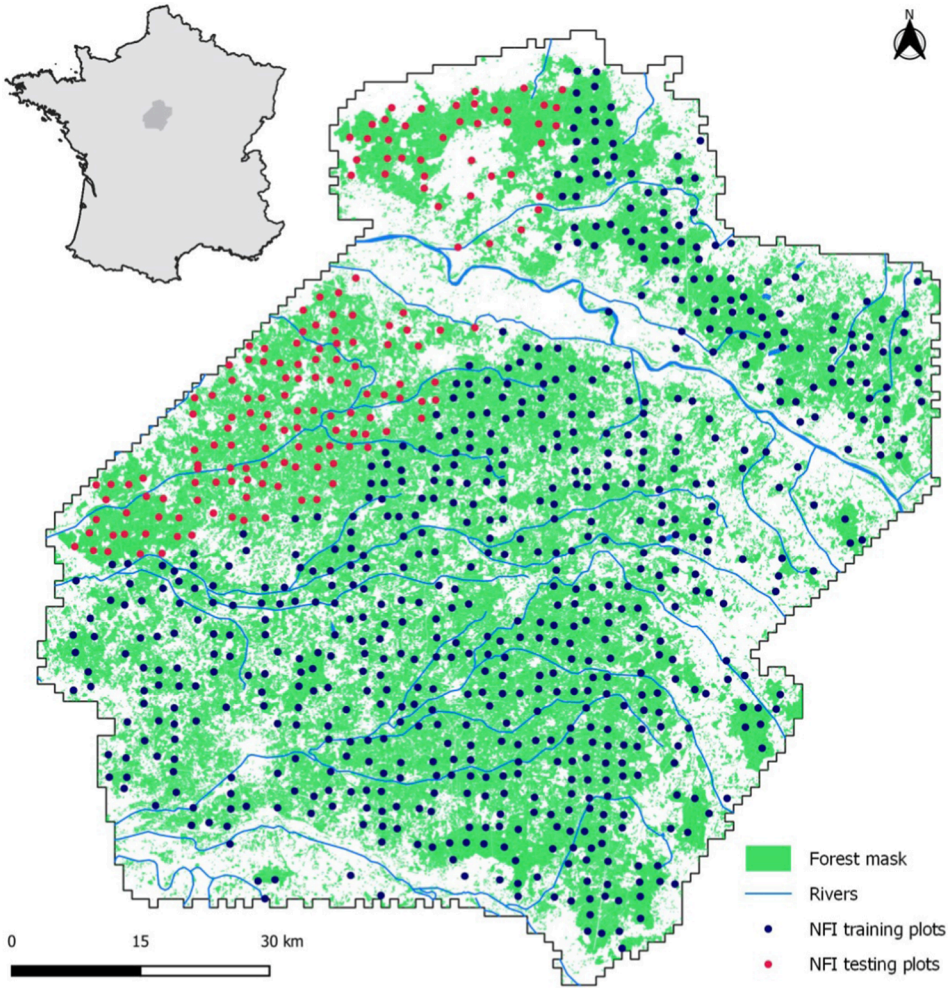

Fig. 1. Location of the study area and spatial distribution of the NFI field plots. The colour of NFI plots represents the training and testing set used for modelling (Section 3.1).

are based on a rigorous statistical framework for designing sampling and inventory procedures with corresponding estimators (Tewari and Kleinn, 2015), and can deliver precise estimates of forest attributes such as growing stock volume (GSV) and basal area (BA) at national down to regional scales (Brosofske et al., 2014; Kangas et al., 2018). Below these scales, the estimation domain is too sparsely sampled, resulting in poor estimation precision of forest attributes (Coelho and Pereira, 2011). Such a decrease in estimation precision prevents forest stakeholders from using NFI estimates to develop strategies and support decision-making processes over small functional-management territories (Vega et al., 2021). To support these activities, NFI precision has to be improved and complemented with map products providing detailed information about the amount and localization of the forest resources (Chirici et al., 2020, Rahlf et al., 2021).

Multisource National Forest Inventory (MSNFI) approaches has made it possible to answer these specific needs and also encourage wall-to-wall mapping of forest attributes at high spatial resolution (Saarela et al., 2020), through a statistical combination of NFI and auxiliary data (Tomppo et al., 2008). The auxiliary datasets such as thematic maps (i.e. forest maps, soil maps) and remote sensing products (i.e. optical, SAR, 3D point cloud data) are selected based on their availability, correlation with the forest attributes of interest and their limited cost.

Many optical remote sensing satellites such as SPOT, Landsat, IKONOS, GeoEye-1, Quickbird, and Sentinel-2 have been tested and used in various forest inventories programmes around the world (Castillo-Santiago et al., 2010; Irulappa-Pillai-Vijayakumar et al., 2019; Puliti et al., 2021). Despite their large spatial coverage and high temporal resolution, the availability of optical imagery is often impeded by cloud cover, haze, cloud shadow, while the derived vegetation indices tends to saturate in mid biomass (~130 Mg/ha) (Zhao et al., 2016). Airborne remote sensing data from airborne laser scanning (ALS) and digital aerial photogrammetry (DAP) (Gobakken et al., 2015; Irulappa-Pillai-Vijayakumar et al., 2019) proved to be the most valuable sources of auxiliary data to characterize the forest canopy structure and estimate the most important forest attributes (Durrieu et al., 2015; White et al., 2013). Unlike ALS, DAP cannot provide forest structure information below the canopy cover, and for the computation of canopy height model, DAP relies on a digital terrain model (DTM) from an external source (St-Onge et al., 2008). After conducting analytical research on 70 peer-reviewed studies, Zolkos et al. (2013) concluded that ALS proved to be performing significantly better for forest resource estimation than any other data source. However, due to the low cost of DAP, the data are updated regularly which makes it appealing to use for forest monitoring (Rahlf et al., 2017). Hence, the debate on which auxiliary data is the most suitable for forest resource estimation in MSNFI remains largely open.

The relationships between NFI and auxiliary data were investigated using various approaches. Numerous parametric (i.e. linear regression, Bayesian regression model, geographically weighted regression) and non-parametric (i.e. random forest, *k*-nearest neighbours or *k*-NN) methods have been implemented for downscaling NFI estimates over small management units (Barrett et al., 2016; Gregoire et al., 2016; Irulappa-Pillai-Vijayakumar et al., 2019; Tomppo et al., 2008). These

Table 1
Mean and standard deviation (SD) of forest attributes, for all plots (819 plots), training (80% of the plots) and testing (20 % of the plots) sets. Training and testing sets were used for model development (Section 3.1).

Forest attribute	All plots		Training plots		Testing plots	
	Mean	SD	Mean	SD	Mean	SD
Growing stock (GSV, m^3ha^{-1})	160.7	111.3	160.0	110.0	162.9	116.7
Basal area (BA, m^3ha^{-1})	21.1	11.8	21.2	11.9	20.4	11.5
Volume increment (VI, $m^3ha^{-1}yr^{-1}$)	5.7	3.9	5.7	3.9	5.3	3.8

methods are already operational in the boreal region (Kangas et al., 2018), while experiences are limited in some temperate regions due to the complexity in forest composition and structure (Vega et al., 2021). Among all these methods, *k*-NN technique gained popularity in MSNFIs (Irulappa-Pillai-Vijayakumar et al., 2019; Magnussen et al., 2009). A strong argument for *k*-NN comes from its ease of implementation, its ability to predict multiple forest attributes simultaneously (Goerndt et al., 2019), and the absence of assumption regarding the distribution of predictor and response variables (Irulappa-Pillai-Vijayakumar et al., 2019).

Similarly, various inference frameworks have been used to estimate population parameters, including total, mean, and variance from plot and auxiliary data (Gregoire et al., 2016). Since NFI relies on sampling procedures, model-assisted inference has been largely used to build upon the properties of the sampling design for estimation (Saarela et al., 2015). Small area estimation technique has been developed to downscale the estimates to small geographical domains such as municipalities and forest districts (Breidenbach and Astrup, 2012; Tomppo et al., 2008). However, a minimum number of sample plots is required in each domain to allow estimation, thus constraining the geographical scale at which forest attributes could be reported with a requested precision (Vega et al., 2021). The model-based inference framework allows more flexibility on the sampling design and could provide estimates even in those domains that do not include any field plot (Ståhl et al., 2016). As the estimates are directly derived from the model, the emphasis is placed on the model specification to minimize bias (Longford, 2005). The downside of *k*-NN technique in this context is the absence of a well-established theoretical inference framework, although several estimators have already been proposed (McRoberts et al., 2011). Furthermore, the incapacity of k-NN to predict outside of the calibration domain (James et al., 2013) requires the development of specific approaches to account for prediction bias (Magnussen et al., 2010). With much less sampled plots than auxiliary data, MSNFIs are naturally put in a situation of imbalance that translates into frequent model extrapolation (Meyer and Pebesma, 2021). This stresses the importance of addressing model extrapolation while analysing and quantifying the reliability of the predicted forest attributes.

Models have also largely been used to produce wall-to-wall maps of forest attributes (Chirici et al., 2020; Coops et al., 2021; Nilsson et al., 2017), and error propagation methods have been proposed to assess prediction uncertainties (Blackard et al., 2008; Esteban et al., 2019; Mascaro et al., 2011). But the reliability of pixels predictions in the context of statistical inference have rarely been appreciated (Saarela et al., 2020).

The objective of this study was to propose a method to adapt the model-based *k*-NN estimation framework of population parameters for the production of high resolution forest attribute and reliability maps at pixel level. We defined prediction reliability at pixel level in terms of precision and accuracy. Precision was derived through resampling, building upon a bootstrap model-based estimator proposed by McRoberts et al. (2022) for population parameters. Prediction accuracy was defined through a measure of extrapolation, to account for *k*-NN's inability to predict outside of its calibration domain (Magnussen et al., 2010). Extrapolation was assessed using an original convex hull approach applied to the auxiliary variable space (Conn et al., 2015; Cook, 1977). The method was evaluated for three major forest attributes, GSV, BA, and growing stock volume increment (VI) in a large territory composed of complex forest structures. The auxiliary data included a forest map, Landsat 8 images and 3D point cloud data from ALS and DAP.

2. Material

2.1. Study site

The study was carried out in the temperate forest of Sologne and Orléans, located in the French region of Centre Val De Loire (Fig. 1). The area covers 7,694 km^2 which extends from 1.41° to 2.79° East Longitude and 47.1° to 48.1° North Latitude. The area is characterized by a semi-oceanic climate, with mean annual precipitation of 731 mm and a mean annual temperature of 10.9 °C. Elevation ranges from 70 to 180 m and topography represents a flat terrain. The soil is predominantly made up of sand and clay, originating from the erosion of Massif Central. Because of the low permeability of these soils, the area has numerous ponds and swamps. About 53% (3,900 km^2) of the area is covered by forests. Forested areas are primarily dominated by broad-leaved stands (74%) of oaks (i.e., *Quercus petraea* Mill. and *Quercus robur* L.). Coniferous stands account for 16% of the forested area and are dominated by scots pines (*Pinus sylvestris* L.) and maritime pines (*Pinus pinaster* Ait.). The remaining forest area is covered by a mixed stand (10%) of oaks and Scots pines. The forest is characterized by different management policies influenced by forest ownership. The northern part of the study area is dominated by state-owned forests which are intensely managed. Private forests represent ~ 84% of the forested area.

2.2. Field inventory data

A total of 819 field plots were used. The plots were surveyed on the field between 2010 and 2014 in the framework of the NFI. Such an aggregation is used to produce official statistics at national and regional level (Robert et al., 2009). The plot sample was drawn in such a way that the last annual sample (2014) corresponded to the reference year for the auxiliary data described in Section 2.3.

Field information is collected on 4 concentric circular plots of radius 6, 9, 15, and 25 m. The largest concentric plot is used to describe the stand structure and composition. Tree measurements are restricted to the first three concentric plots according to the tree circumference at breast height (1.3 m). Small ([23.5–70.5 cm[), medium ([70.5–117.5 cm [), and large ([117.5 cm[) trees are measured in plots of radius 6, 9, and 15 m respectively. The information about species, vegetation state, 5-year radial increment, and circumference at breast height are collected for all the trees fulfilling the circumference constraints. Tree attributes such as stem volume and stem volume increment are computed from field measurements using NFI's species-specific allometric models. Plot position on the field is fixed using low-cost GPS, interpretation of printed aerial photographs, and/or routing and chaining. The precision of the plot centre position is evaluated at 3.67 m (±3.3 m) (Guitet, 2018).

Plot-level forest attributes are computed from the tree measurements and their inclusion probability, which is based on the tree position within the plot (Hervé et al., 2014). The study focuses on three forest attributes of importance: GSV (m^3ha^{-1}), computed as the weighted sum of the individual tree stem volume (stem diameter up to 7 cm), BA (m^3ha^{-1}), defined as the weighted sum of the tree cross-sectional area at breast height, and VI ($m^3ha^{-1}yr^{-1}$), computed as the weighted sum of the mean annual volume increment over 5 years. Summary statistics of the three forest attributes are shown in Table 1.

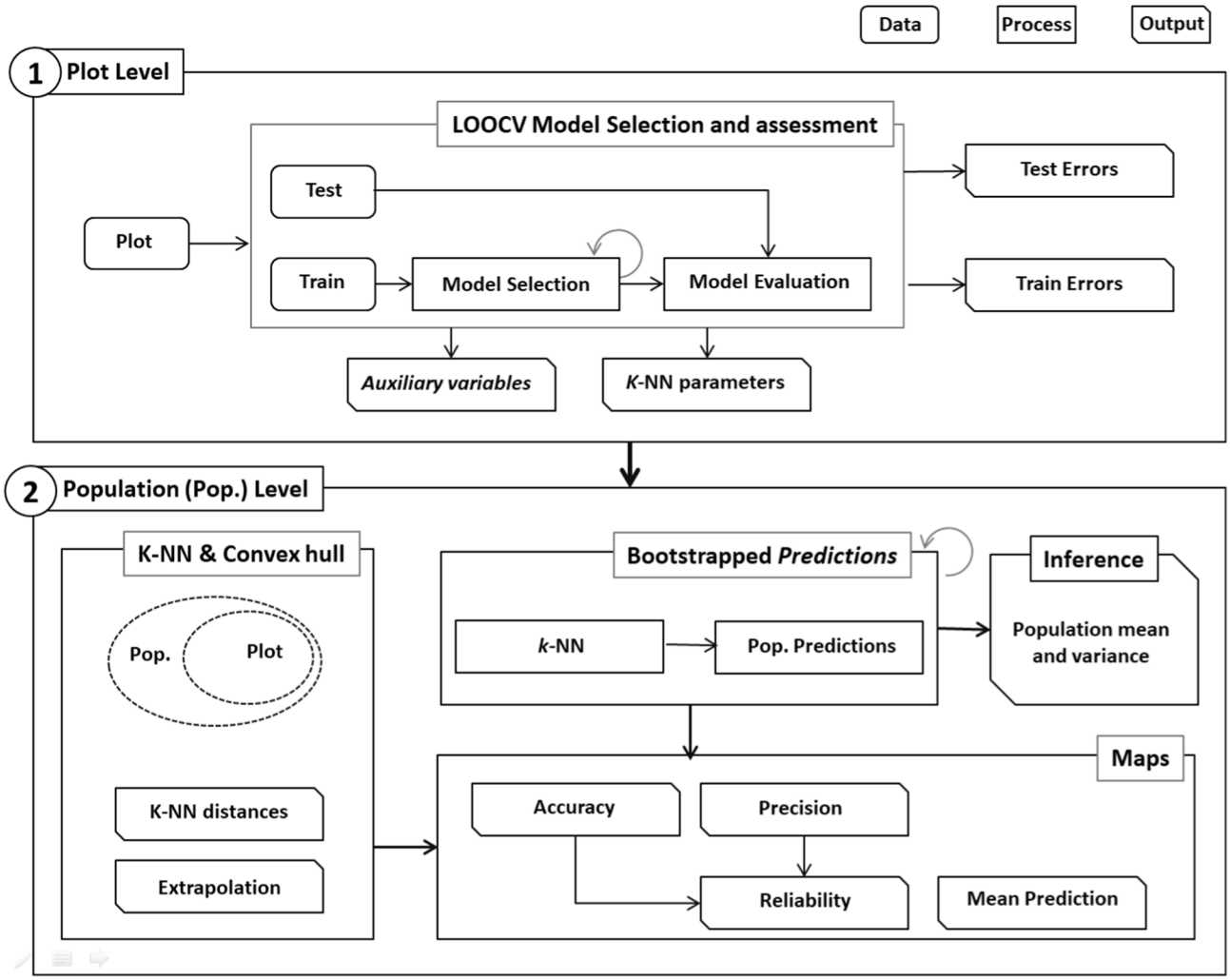

Fig. 2. Overview of the method, including (1) plot-level model selection and assessment, (2) population inference and maps.

2.3. Auxiliary data and processing

Three freely available auxiliary data were considered in this study: forest map, Landsat images, and 3D point clouds. Two sources of point cloud were used for comparison: ALS, and DAP.

The auxiliary data were processed and integrated into a grid of 30 m resolution, corresponding to the diameter of the NFI field plots. Each pixel of the grid had a surface of 900 m^2.The origin of the grid was set according to the Lambert conic projection using Lambert-93 geodetic system (EPSG: 2154). Data processing and analysis were conducted in R (Version 4.0.2) environment, using lidR (Roussel et al., 2020), yaImpute (Crookston and Finley, 2008), and geometry (Habel et al., 2019) packages.

2.3.1. Forest map

The forest map (BD-Forêt® version 2.0) is an open-source vector map, defining the geographical reference for the forest land cover and the description of forest stands. The map is produced by the National Institute of Geographic and Forest Information (IGN), through semi-automatic segmentation of near-infrared aerial images (BD ORTHO®) with a resolution of 0.5 m, followed by the photo-interpretation of the resulting segments. The map provides information about forest area and composition for the dominant species within a forest patch of at least 0.5 ha.

The forest map was used to generate a forest mask to limit the spatial extent of the pixels within the forest land cover only. Each pixel was characterized by a dominant forest type (pure conifers, pure broad-leaved, or mixed stands), which was achieved by selecting the forest type at the pixel centre (Irulappa-Pillai-Vijayakumar et al., 2019).

2.3.2. Landsat data

Four processed Landsat 8 images acquired on 8 September 2014 were used in this study. The data were acquired freely from the Theia portal (*https://theia-landsat.cnes.fr*). The images were processed to level 2A, which includes: orthorectification, cloud and their shadow masking, and atmospheric correction (Hagolle et al., 2015). Landsat 8 Operational Land Imager images consist of nine spectral bands of 30 m spatial resolution. Along with the spectral bands, 12 spectral indices were computed: Normalized Difference Vegetation Index, Normalized Difference Moisture Index, Normalized Difference Water Index, Green Normalized Difference Vegetation Index, Soil Adjusted Vegetation Index, Specific Leaf Area Vegetation Index, Modified Soil Adjusted Vegetation Index, Enhanced Vegetation Index, Simple Ratio, Brightness, Greenness and Wetness. The details for the computation of indices were provided in Irulappa-Pillai-Vijayakumar et al. (2019).

2.3.3. Airborne laser scanning and digital aerial photogrammetry

The Airborne Laser Scanning (ALS) data were collected by IGN, between January and March 2014. The northern part of the study area (2,758 km^2) was covered using Optech ALTM 3100 EA laser scanner operated at an altitude of 1,530 m above ground level (AGL). The pulse repetition frequency was 71 Hz, and the scanning frequency was 35 Hz

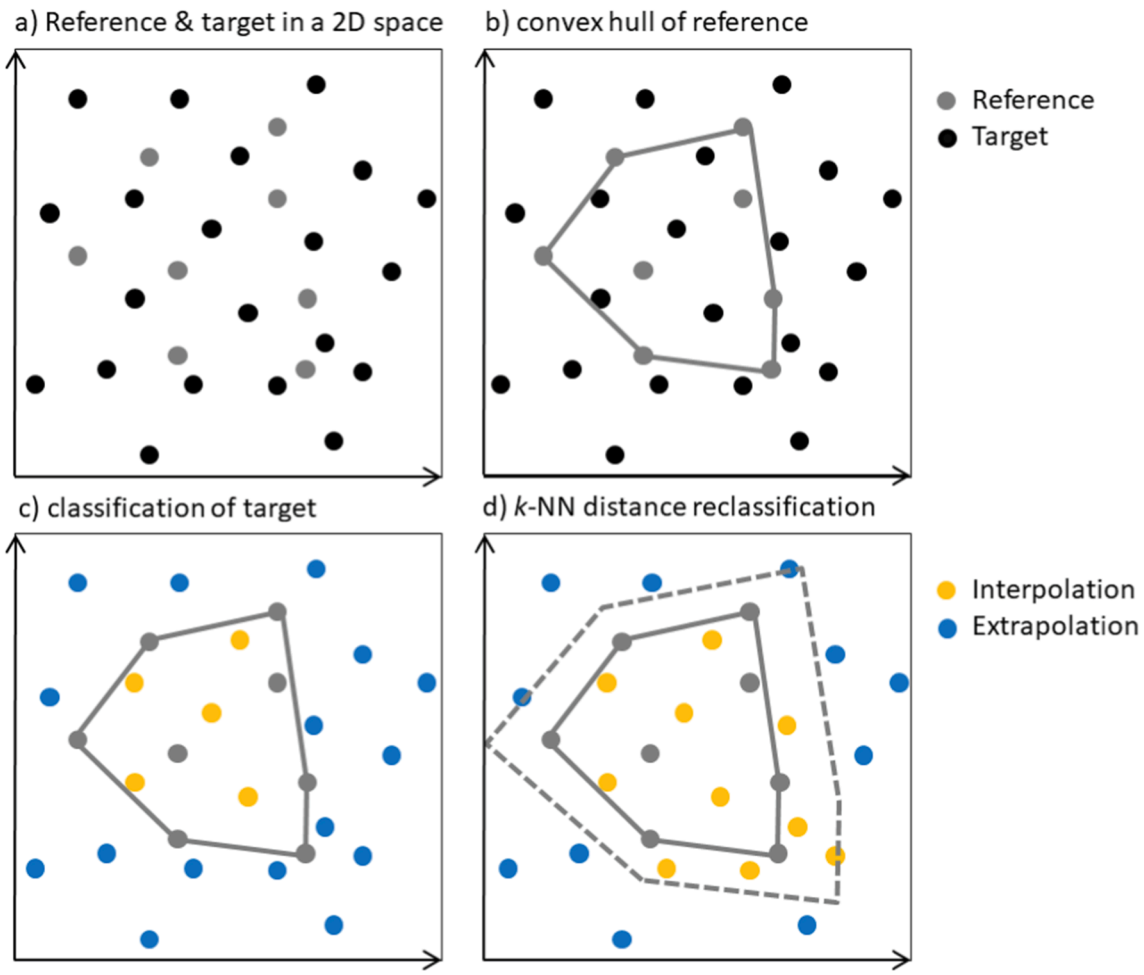

Fig. 3. Scheme of the process of identification and characterisation of extrapolation in a 2D auxiliary variable space: (a) reference (plots) and target (pixels) sets; (b) computation of the convex hull of reference; (c) classification of target into interpolation or extrapolation; (d) reclassification of extrapolated targets using an interpolation distance threshold.

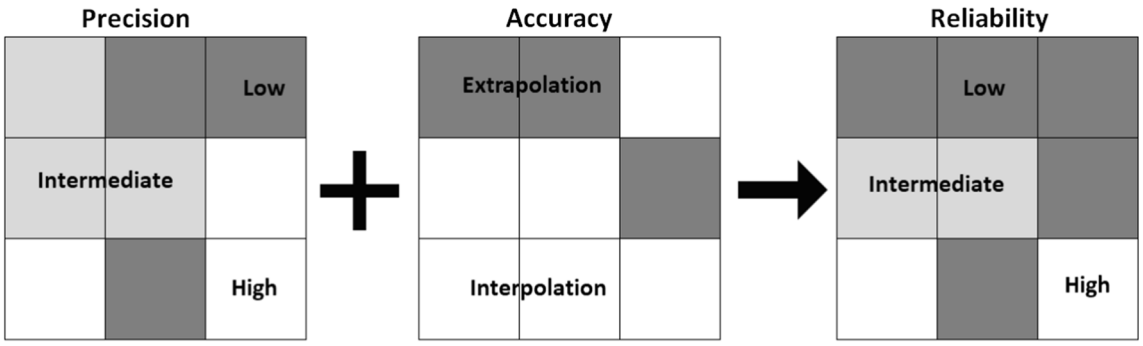

Fig. 4. Scheme of the reliability assessment for a sample of pixel's predictions based on precision and accuracy.

with a maximum scan angle of 25°. A supplementary flight was conducted to cover the southern part of the study site (4,936 km^2). The data were acquired using a Leica ALS70 sensor flying at an altitude of 1,797 m AGL. The pulse repetition frequency was 309 Hz, and the scanning frequency was 35 Hz with a maximum scan angle of 25°. The point density of both flights was 2 pts m^{-2}. The point clouds were processed using IGN processing lines using Terrasolid's Terrascan software (*https://terrasolid.com*). The data were delivered as classified point cloud tiles of 1 km resolution. The ground classified points were used to generate a DTM at a 1 m resolution using Triangulated Irregular Network algorithm. The points above the ground level were normalized using the DTM.

The Digital Aerial Photogrammetry (DAP) data were acquired between 2013 (July 06 – July 12) and 2014 (May 18 – July 16), using the IGN CamV2 system (14,650 × 10,700 pixels of 6.8 μm each) (Souchon et al., 2010) mounted with a Zeiss lens of 125 mm focal length. The images were acquired at 6,400 m AGL with 60% overlap along track and 25% across track. The pixel resolution at ground level was 0.35 m. The oriented images were processed using IGN MicMac photogrammetric software (Rupnik et al., 2017). The dense image matching was implemented using a Per Image Matching approach. Depth maps of each image pair were generated at a pyramid level 2, merged, and then

Table 2
Parameters of the selected *k*-NN models for ALS and DAP.

Dataset	*k*-NN models		Selected variables		
	Variable selection	Distance	Forest map	3D point cloud	Landsat
ALS	Deletion	msn	Forest type	50th percentile of heights 4th cumulative percentage of returns Canopy closure ratio	Green band
DAP	Addition	msn	Forest type	Standard deviation 40th percentile of heights	Green Short-wave infrared 1

converted into a 3D point cloud. The resulting points (0.9 pts m^{-2}) were further normalized using the ALS DTM.

Multiple metrics were computed from both ALS and DAP point cloud at the grid level. These include: minimum, maximum and mean height, standard deviation, skewness, kurtosis and entropy of height distribution, percentage of returns above mean height, percentile of height distribution, and cumulative percentage of return. The above metrics were computed for all the points above 2 m. ALS metrics were computed for the first and last returns, as well as for the intensity of first returns. Additional metrics were computed for the ratio of area and canopy closure ratio (Rcc: ratio between canopy and ground cover) (Roussel et al., 2020).

3. Methods

The methods included two steps, *(i)* model selection and assessment, and *(ii)* statistical inference and mapping (Fig. 2). The relationships between field attributes and the auxiliary data were investigated using non-parametric multivariate *k*-NN models (McRoberts, 2012).

3.1. Model selection and assessment

Model selection and assessment was achieved using a training–testing approach. The data split was based on a geographical partitioning of the study area (Meyer and Pebesma, 2021). First, the bounding box of the field plots was computed and one of the corners was randomly selected. Then, the 20% of the nearest plots from that corner were selected and retained as the testing set. The training set was formed with the remaining 80% of the plots and was used for model selection and evaluation. As the number of auxiliary variables were high and could be affected by autocorrelation issues, the variables were first thinned using a variable reduction approach.

3.1.1. Variable reduction

The *k*-NN predictions are sensitive to the model dimensionality (Breidenbach et al., 2012). A model with a large number of predictor variables would increase the modelling complexity, computation time, and may degrade its performance (Bhanu and Lin, 2003). Furthermore, some auxiliary variables are auto-correlated (e.g., ALS height percentile), and may cause model overfitting problems (Moser et al., 2017).

Highly correlated auxiliary variables were circumvented through a variable reduction approach. The correlation matrix was computed between auxiliary variables and the targeted field attributes. For each field attribute, the auxiliary variables with the highest correlation were selected to initiate a vector of variables. The remaining auxiliary variables were then processed based on their correlation rank (from higher to lower). An auxiliary variable was added to the vector if its correlation with all the vector variables was lower than a threshold (0.7 units) (Dormann et al., 2013). The auxiliary variable common to all the vectors (one vector per field attribute) were then selected to initiate a vector of reduced variables and their average correlation with field attributes was computed. The remaining auxiliary variables were further ranked according to their mean correlation with the field attributes and tested for inclusion in the vector of reduced variables using the same procedure and parameters.

3.1.2. k-NN models

Along with the number of variables, *k*-NN model parameters include the number of neighbours, a distance metric and a distance weighting function (McRoberts, 2012). The optimal value of *k* is usually a trade-off between bias and variance (McRoberts et al., 2015). Following previous researches, models were implemented using a *k* value of 5 (Irulappa-Pillai-Vijayakumar et al., 2019; Chirici et al., 2020). The *k*-NN were tuned for 2 distances metrics and were built using 1 up to 10 auxiliary variables selected using an iterative feature selection algorithm. For both ALS and DAP, the best model was selected from a total of 40 competing models.

Distance metrics included Euclidean (Eq. (1)) and most similar neighbours (Eq. (2)) (Crookston and Finley, 2008; Moeur and Stage, 1995). Both distances were computed in a normalized variable space by dividing their mean-centred values with their standard deviation and using an inverse distance weighting function (Eq. (3)):

$$d_{ij}(\mathrm{euc}) = \sqrt{(x_i - x_j)^{\mathrm{T}}(x_i - x_j)} \quad (1)$$

$$d_{ij}(\mathrm{msn}) = \sqrt{(x_i - x_j)\Gamma\Lambda^2\Gamma^{\mathrm{T}}(x_i - x_j)^{T}} \quad (2)$$

Table 4
Test model prediction means and errors for ALS and DAP based predictions of growing stock volume (GSV, m^3ha^{-1}), basal area (BA, m^3ha^{-1}), and volume increment (VI, $m^3ha^{-1}yr^{-1}$). Numbers in parenthesis are relative errors in percent.

Dataset	Forest attribute	Prediction Mean	RMSE	ME
ALS	GSV	158.1	80.4 (49.3)	−4.8 (−3.0)
	BA	20.4	9.4 (46.2)	0.1 (0.3)
	VI	5.2	3.6 (67.4)	−0.1 (−2.2)
DAP	GSV	159.7	79.6 (48.8)	−3.5 (−2.2)
	BA	21.0	9.6 (47.2)	0.6 (3.2)
	VI	5.6	3.4 (64.2)	0.3 (4.9)

Table 3
Cross-validation model prediction means and errors for ALS and DAP based prediction of growing stock volume (GSV, m^3ha^{-1}), basal area (BA, m^3ha^{-1}), and volume increment (VI, $m^3ha^{-1}yr^{-1}$). Numbers in parenthesis are relative errors in percent.

Dataset	Forest attribute	Calibration			Validation		
		Prediction Mean	RMSE	ME	Prediction Mean	RMSE	ME
ALS	GSV	159.8	59.3 (37.1)	−0.1 (−0.1)	159.6	76.4 (47.8)	−0.4 (−0.2)
	BA	21.2	7.6 (35.9)	0.1 (0.3)	21.2	9.7 (45.7)	0.1 (0.3)
	VI	5.7	2.9 (50.0)	< 0.1 (0.2)	5.7	3.6 (63.4)	< 0.1 (0.1)
DAP	GSV	159.2	59 (36.9)	−0.6 (−0.4)	159.6	75.2 (47.0)	−0.3 (−0.2)
	BA	21.2	7.6 (35.7)	< 0.1 (<0.1)	21.3	9.5 (45.0)	0.1 (0.5)
	VI	5.7	2.8 (48.2)	−0.1 (−1.0)	5.7	3.6 (61.9)	> −0.1 (−0.7)

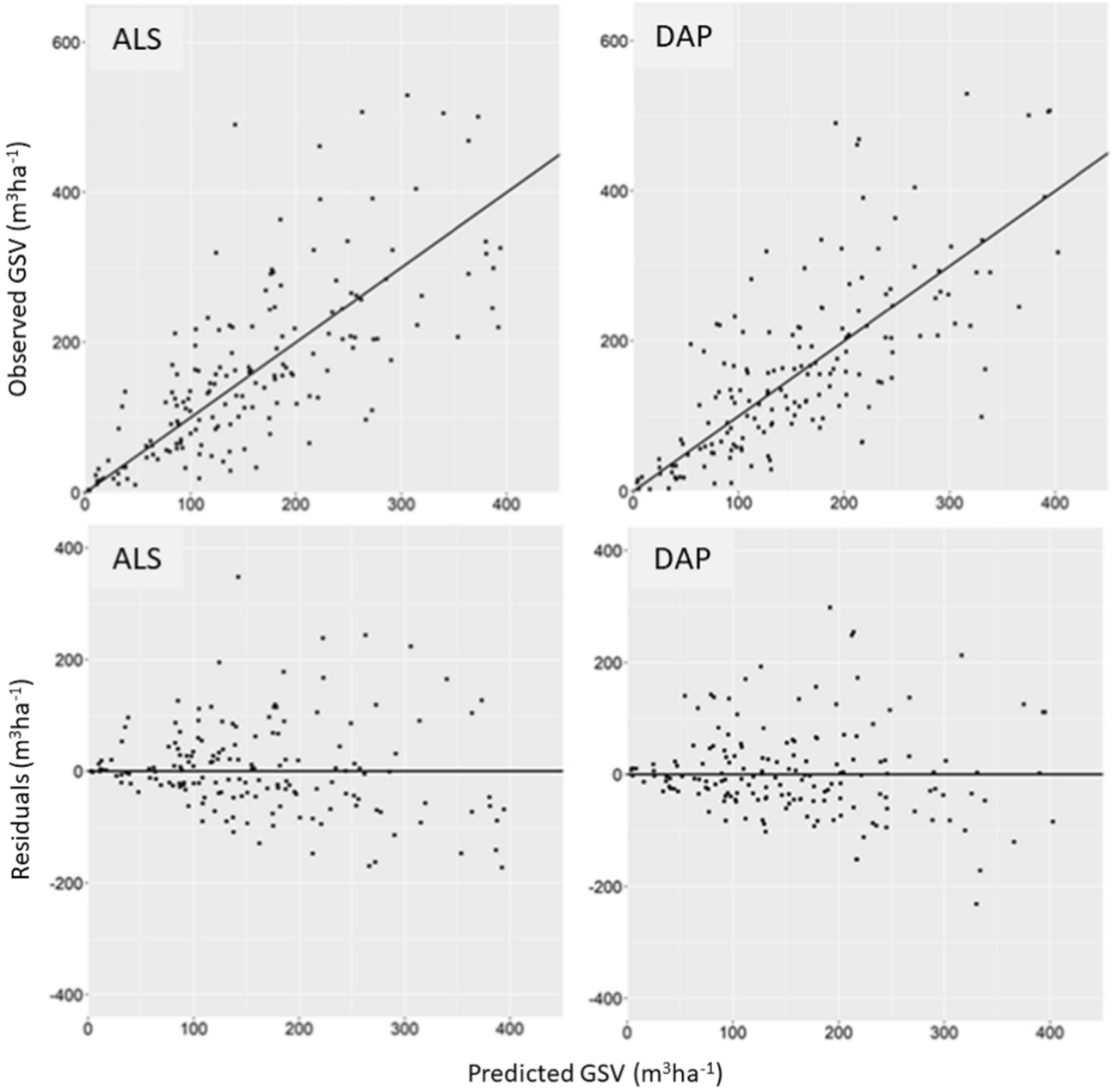

Fig. 5. Graphs of the observed versus predicted growing stock volume (GSV) and of corresponding prediction residuals over the test set for both ALS and DAP based models.

Table 5
Model-based bootstrapped population mean ($\hat{\mu}_{boot}$), variance ($\hat{V}_{boot}$), and bias ($\hat{B}_{boot}$) of predicted forest attributes with ALS and DAP based models.

Dataset	Forest attribute	$\hat{\mu}_{boot}$	$\hat{V}_{boot}(\hat{\mu})$	$\hat{B}_{boot}(\hat{\mu})$
ALS	GSV (m^3ha^{-1})	161.9	6.4	0.2
	BA (m^3ha^{-1})	20.8	0.1	< 0.1
	VI ($m^3ha^{-1}yr^{-1}$)	5.6	< 0.1	< 0.1
DAP	GSV (m^3ha^{-1})	160.7	5.1	0.1
	BA (m^3ha^{-1})	20.7	0.1	< 0.1
	VI ($m^3ha^{-1}yr^{-1}$)	5.5	< 0.1	< 0.1

$$w_{ij} = \frac{1}{1 + d_{ij}} \quad (3)$$

where, $d_{ij}(euc)$ is the Euclidean distance, $d_{ij}(msn)$ is the most similar neighbours distance, x_i is the vector of target auxiliary variables, x_j is the vector of reference auxiliary variables, Λ is the matrix of canonical correlations, Γ is the matrix of canonical vectors of auxiliary data resulting from the canonical correlation between auxiliary data and field attributes, T denotes the transpose of the matrix, and w_{ij} is the weight applied to d_{ij} to compute predictions. While $d_{ij}(euc)$ only relies on auxiliary variables, $d_{ij}(msn)$ also preserves the covariance structure of the field attributes.

Variable selection was performed using an iterative feature selection algorithm applied in both addition and deletion modes (Crookston and Finley, 2008). The algorithm was developed for *k*-NN models and relied on the minimization of the Root Mean Square Error (RMSE, Eq. (4)) of

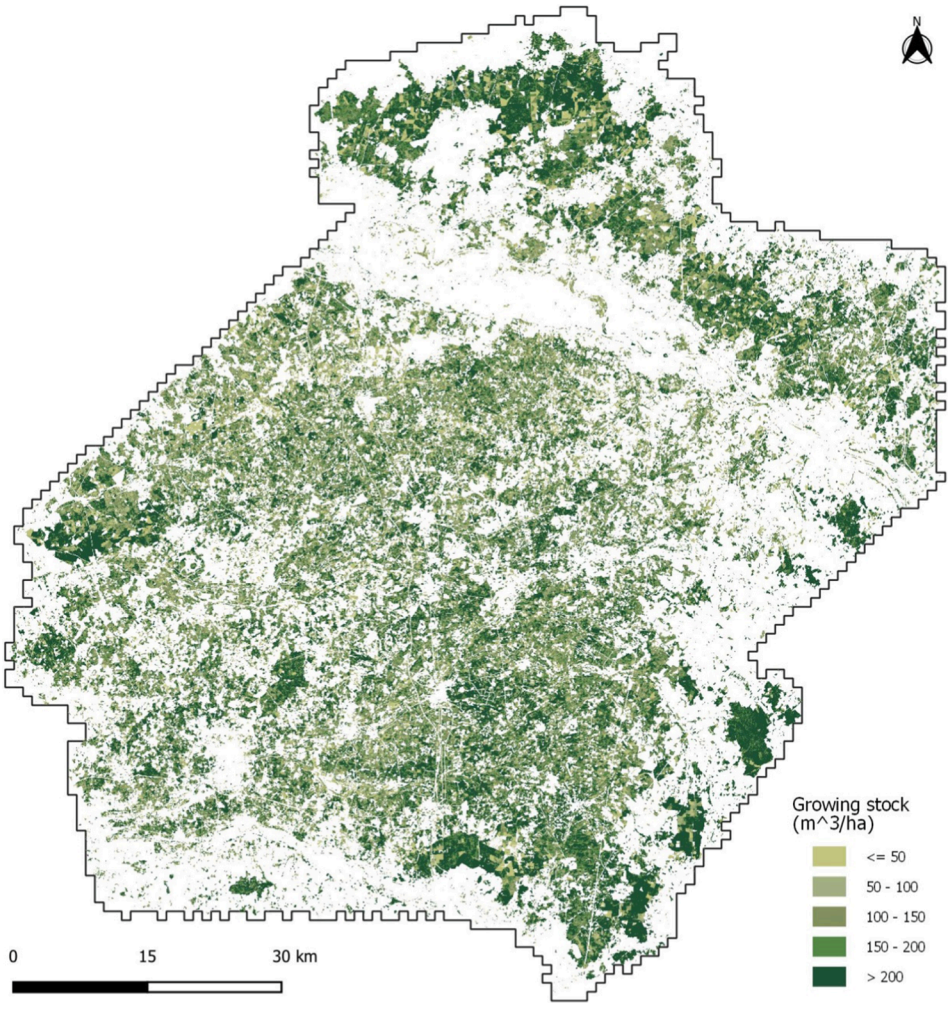

Fig. 6. Map of ALS growing stock volume predictions at 30 m spatial resolution.

predictions. The algorithm was applied independently for each distance metric, and for each case, the best 10 variables were selected for further processing.

The nested models made of 1 up to 10 variables were evaluated using leave-one-out cross-validation. For each model, the cross-validated RMSE (Eq. (4)), mean error (ME, Eq. (5)) and total error (TE, Eq. (6)) (Hou et al., 2017; Vega et al., 2021) were computed for both calibration and validation sets based on the following equations:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}; \quad \mathrm{rRMSE} = \frac{\mathrm{RMSE}}{\bar{y}} \times 100 \tag{4}$$

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i); \quad \mathrm{rME} = \frac{\mathrm{ME}}{\bar{y}} \times 100 \tag{5}$$

$$\mathrm{TE} = \sqrt{\mathrm{RMSE}^2 + \mathrm{ME}^2}; \quad \mathrm{rTE} = \frac{\mathrm{TE}}{\bar{y}} \times 100 \tag{6}$$

where y_i is the observed value, $\hat{y}_i$ is predicted value, $\bar{y}$ is the mean of observed values and n is the number of NFI plots.

The optimal model was selected for both ALS and DAP in two steps. First, models exhibiting overfitting were filtered out from the selection. Overfitting was identified using the method proposed by Valbuena et al. (2017). It consisted in computing the ratio between the square error achieved on the cross-validation dataset and the one achieved using the entire train dataset (SSR, Eq. (7)), and rejecting models having a ratio value above 1.1:

$$\mathrm{SSR} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{cv} - \hat{y}_i\right) \Big/ \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{train} - \hat{y}_i\right) \tag{7}$$

where SSR is the overfitting index, y_i^{cv} are the cross-validated predictions and y_i^{train} are the training predictions.

From the set of remaining models, the models having a rTE within 1% of the lowest rTE were selected and the best model was defined as the one having the lowest number of predictors for model parsimony.

Finally, a k-NN was built using the whole train set and evaluated on the test set as an independent performance assessment. Performance measures included both RMSE and ME.

3.2. Statistical inference and mapping

Predictions at the population level were obtained using bootstrap (N = 100). The size of the bootstrap sample was set to the number of field plots. *k*-NN models were trained for each bootstrapped sample and used to predict the entire population of pixels, resulting in 100 predictions for each pixel and forest attribute.

3.2.1. Statistical inference

The population mean (Eq. (8)) and variance (Eq. (9)) from the field plots were computed using a simple expansion estimator (Magnussen et al., 2020), and were used to evaluate the *k*-NN performance at the population level:

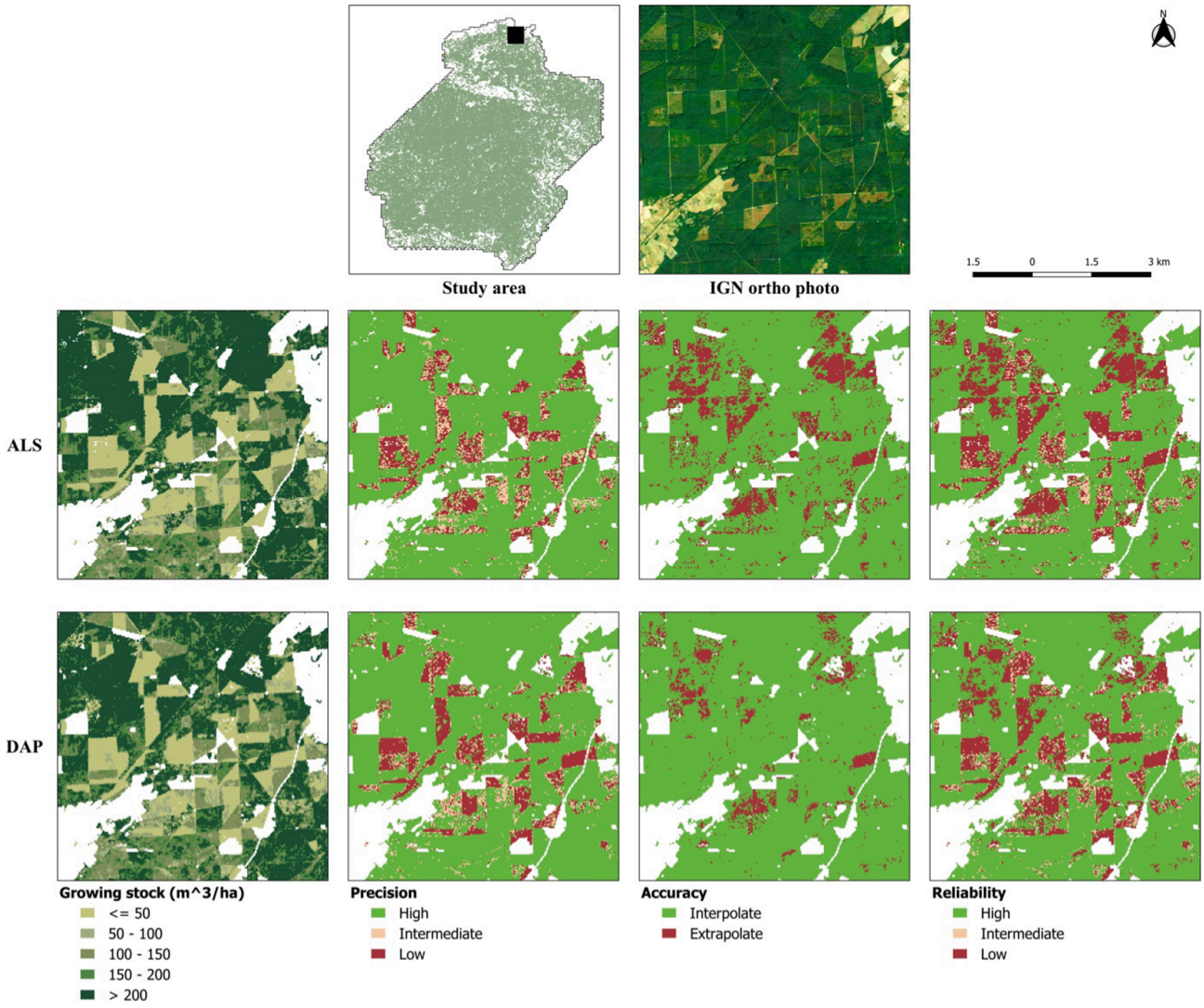

Fig. 7. Maps of predicted growing stock volume, and corresponding precision, accuracy and reliability, for both ALS and DAP based models.

Table 6
Mean predictions, mean *k*-NN distances and proportions of pixels per precision classes (H: high; I: intermediate; L: low) for both ALS and DAP-based models. The unit of the mean estimates corresponds to the unit of the forest attributes: growing stock volume (GSV, m^3ha^{-1}), basal area (BA, m^3ha^{-1}) and volume increment (VI, $m^3ha^{-1}yr^{-1}$). *k*-NN distances are unit less.

Dataset	Attribute	Mean predictions			Mean distance			Pixels (%)		
		H	I	L	H	I	L	H	I	L
ALS	GSV	176.2	47.0	21.0	0.13	0.21	0.25	89.9	5.3	4.8
	BA	22.0	7.0	5.0	0.13	0.21	0.26	92.0	5.4	2.6
	VI	5.7	4.1	1.5	0.13	0.24	0.29	97.2	2.7	0.2
DAP	GSV	179.4	54.5	24.3	0.11	0.15	0.15	86.4	7.8	5.8
	BA	22.3	8.9	4.9	0.11	0.15	0.16	89.6	6.5	3.9
	VI	5.7	2.1	1.6	0.11	0.16	0.20	95.7	3.7	0.5

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{8}$$

$$\widehat{V}(\widehat{\mu}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(y_i - \widehat{\mu})^2 \tag{9}$$

where n is the number of sample plots and y_i is the plot observed value.

The bootstrapped model-based estimator was used to estimate the population parameters from *k*-NN predictions (McRoberts et al., 2022). The mean (Eq. (10)), bias (Eq. (11)), and variance (Eq. (12)) estimators were computed as follow:

$$\widehat{\mu}_{boot} = \frac{1}{n_{boot}}\sum_{b=1}^{n_{boot}} \widehat{\mu}^{b}_{boot} \tag{10}$$

$$\widehat{B}_{boot}(\widehat{\mu}) = \widehat{\mu}_{boot} - \widehat{\mu} \tag{11}$$

$$\widehat{V}_{boot}(\widehat{\mu}) = \frac{1}{n_{boot}-1}\sum_{b=1}^{n_{boot}}\left(\widehat{\mu}^{b}_{boot} - \widehat{\mu}_{boot}\right)^2 \tag{12}$$

where n_{boot} is the number of bootstrap samples, $\widehat{\mu}^{b}_{boot}$ is the mean of the boot b, $\widehat{\mu}_{boot}$ is the mean of the bootstrap samples.

Table 7
Mean predictions, mean *k*-NN distances and proportions of pixels per accuracy classes (H: interpolation; L: extrapolation) for both ALS and DAP-based models. The unit of the mean estimates corresponds to the unit of the forest attributes: growing stock volume (GSV, m^3ha^{-1}), basal area (BA, m^3ha^{-1}) and volume increment (VI, $m^3ha^{-1}yr^{-1}$). *k*-NN distances are unit less.

Dataset	Attribute	Mean predictions		Mean distance		Pixels (%)	
		H	L	H	L	H	L
ALS	GSV	163.2	143	0.12	0.37	93.7	6.3
	BA	21.1	15.8				
	VI	5.7	4.3				
DAP	GSV	159.8	178.8	0.10	0.37	95.7	4.3
	BA	20.8	19.9				
	VI	5.6	5.5				

3.2.2. Map products

The *k*-NN prediction maps were generated for each forest attribute at the grid pixel resolution (30 m) by computing the mean of the bootstrapped predictions. Along with the prediction maps, three additional maps per forest attribute were produced for the assessment of precision, accuracy and reliability.

A map of precision was generated based on the bootstrapped pixel-level predictions. The precision was assessed through the calibration–validation model rRMSEs (Eq. (4)) and coefficient of variation (CV_{boot}) (Eq. (13)) of the bootstrapped prediction at pixel level.

$$CV_{boot} = \sum_{i=1}^{n_{boot}} \left(\sqrt{\frac{(x_i - \bar{x})^2}{n_{boot} - 1}} \Big/ \bar{y} \right) \Big/ n_{boot} \quad (13)$$

where, n_{boot} is the number of bootstrap samples, x_i and $\bar{x}_i$ are respectively the prediction for a given sample and the prediction mean over the entire samples. For each pixel, the precision was considered high when CV_{boot} was below calibration rRMSE, intermediate when CV_{boot} was in between calibration and validation rRMSE, and low otherwise. Such a qualitative representation was preferred over a quantitative map for operational reasons.

A map of accuracy was computed from the extrapolation status of the pixels. Since *k*-NN could not extrapolate, accurate predictions could only be expected from interpolated values. Extrapolation was identified in the auxiliary data space, using a convex hull approach (Fig. 3) (Barber et al., 1996). The convex hull was built using the auxiliary space describing the field plots. Pixels were then classified as interpolation or extrapolation according to their position with respect to the convex hull (Fig. 3c). The impact of extrapolation on prediction accuracy is expected to be a function of the extrapolation distance. Thus, extrapolated pixels were further classified according to a distance threshold. This threshold was defined as the mean *k*-NN distance of interpolated pixels plus two times its standard deviation. Extrapolated pixels having a *k*-NN distance lower than this threshold were reclassified as interpolated. Finally, the accuracy of pixels was considered high when pixels were classified interpolated and low otherwise.

The third map relates to the reliability of the pixel predictions. From precision and accuracy maps, predictions were considered unreliable irrespective of their precision class when they were found as extrapolated. Otherwise, the reliability was set in accordance with the precision classes (Fig. 4).

4. Results

4.1. Model development and assessment

The selected *k*-NN models are presented in Table 2. For ALS, 5 variables were selected using a variable deletion algorithm and the msn distance. The selected variables are the forest type, the Landsat green band and 3 ALS metrics describing the canopy height (i.e. the 50th percentile of the height distribution, corresponding to the median height), the canopy density (i.e., the 4th cumulative percentage of return computed as the proportion of returns above the 4 ^{h} decile of the height distributions to the total number of returns) and canopy opening (canopy closure ratio). For DAP, the selected model also included 5 variables and was generated using the variable addition algorithm and msn distance. Like ALS, the forest type and the Landsat green band contributed to the model. The 3 remaining auxiliary variables were the short-wave infrared band 1 (SWIR1) from Landsat, and 2 point cloud metrics describing the vertical structure of the canopy (i.e. the standard deviation and the 40th percentile of height distribution).

Table 3 and 4, and Fig. 5 show the performance of the model with

Table A1
Test model prediction means and errors for inverted ALS and DAP based predictions of growing stock volume (GSV, m^3ha^{-1}), basal area (BA, m^3ha^{-1}), and volume increment (VI, $m^3ha^{-1}yr^{-1}$). Numbers in parenthesis are relative errors in percent. Inverted models were built using auxiliary data selected for ALS to DAP and conversely.

Dataset	Forest attribute	Prediction Mean	RMSE	ME
ALS	GSV	163.8	84.4 (51.8)	0.8 (0.5)
	BA	21.3	10.1 (49.4)	0.9 (4.6)
	VI	5.7	3.7 (69.2)	0.4 (7.2)
DAP	GSV	163.8	80.5 (49.4)	0.6 (0.4)
	BA	21.4	9.7 (47.4)	1 (4.8)
	VI	5.7	3.7 (69.4)	0.4 (8.4)

Table A2
Model-based bootstrapped population mean ($\hat{\mu}_{boot}$), variance ($\hat{V}_{boot}$), bias ($\hat{B}_{boot}$), of predicted forest attributes with inverted ALS and DAP based models. Inverted models were built using auxiliary data selected for ALS to DAP and conversely.

Dataset	Forest attribute	$\hat{\mu}_{boot}$	$\hat{V}_{boot}(\hat{\mu})$	$\hat{B}_{boot}(\hat{\mu})$
ALS	GSV (m^3ha^{-1})	162.3	7.2	0.6
	BA (m^3ha^{-1})	20.9	0.1	0.1
	VI ($m^3ha^{-1}yr^{-1}$)	5.6	< 0.0	< 0.0
DAP	GSV (m^3ha^{-1})	161.8	5.3	−0.1
	BA (m^3ha^{-1})	20.8	0.1	< 0.0
	VI ($m^3ha^{-1}yr^{-1}$)	5.5	< 0.0	< 0.0

Table 8
Mean predictions, mean *k*-NN distances and proportions of pixels per reliability classes (H: high; I: intermediate; L: low) for both ALS and DAP-based models. The unit of the mean estimates corresponds to the unit of the forest attributes: growing stock volume (GSV, m^3ha^{-1}), basal area (BA, m^3ha^{-1}) and volume increment (VI, $m^3ha^{-1}yr^{-1}$). *k*-NN distances are unit less.

Dataset	Attribute	Mean predictions			Mean distance			Pixels (%)		
		H	I	L	H	I	L	H	I	L
ALS	GSV	174.0	48.8	103.1	0.12	0.17	0.31	86.3	4.3	9.4
	BA	22.0	7.6	13.6	0.12	0.17	0.34	87.9	4.2	7.9
	VI	5.7	4.0	4.2	0.12	0.19	0.37	91.7	1.9	6.4
DAP	GSV	177.2	53.7	93.8	0.1	0.12	0.24	83.5	7.0	9.5
	BA	22.2	8.8	13.2	0.1	0.13	0.26	86.4	5.8	7.7
	VI	5.7	2.0	5.1	0.1	0.13	0.35	92.0	3.3	4.7

training and testing data. For the ALS-based model, the cross-validated rRMSEs ranged from 35.9% for BA to 50% for VI in calibration and from 45.7% for BA to 63.4% for VI in validation. The rME was within 0.3% in both calibration and validation for the three forest attributes surveyed (Table 3). The results achieved for the test set are of the same magnitude as the validation set. The test rRMSEs ranged from 46.2% to 67.4% for BA and VI respectively (Table 4). The lowest rME was −3 % for GSV.

For the DAP-based model, the cross-validated rRMSEs ranged from 35.7% for BA to 48.2% for VI in calibration and from 45.0% to 61.9% for the same attributes in validation. The rME were slightly greater than those obtained for ALS. The calibration rME ranged from −1.0% for VI to<0.1% for BA. The validation rME was of the same magnitude, with a minimum value of −0.7% for VI and a maximum of 0.5% for BA (Table 3). The testing rRMSEs varied from 47.2% to 64.2% for BA and VI respectively. The rME was negative for GSV (−2.2%) and positive for both BA and VI, with the greatest value achieved for VI (4.9%) (Table 4).

4.2. Statistical inference and mapping

Population estimates and corresponding maps are presented in Table 5 and Figs. 6 and 7. The mean GSV was estimated to 161.9 m^3ha^{-1} and 160.7 m^3ha^{-1} with ALS and DAP respectively (Table 5). Corresponding variances were of 6.4 m^3ha^{-1} for ALS and 5.1 m^3ha^{-1} for DAP. Estimates for BA and VI were of the same order with both models. The largest volumes are predominantly found in public forests, localized in the northern, south-eastern and western part of the study area (Fig. 6).

Table 6 showed the distribution of predictions within the three precision classes (see Section 3.2.2) for the 3 surveyed forest attributes. For both models, the mean attribute estimations decreased through the precision classes, indicating that lower precisions are associated to the lower attribute values. The mean *k*-NN distances followed the inverse trend, showing the lesser values in the high precision class and the greater values in the low precision class. Overall, the high precision class account for more than 86% of the pixels, ranging from 89.9% (GSV) to 97.2% (VI) for ALS and from 86.4% (GSV) to 95.7 (VI) pixels with DAP-based model.

Table 7 showed the ventilation of predictions in accuracy classes. The high accuracy represented respectively 93.7% and 95.7% of the predictions for ALS and DAP. A major difference between accuracy classes was the mean *k*-NN distance. Low accuracies, which corresponded to extrapolation, exhibited mean *k*-NN distances at least 3 times superior to the high accuracy. Mean estimates tended to be greater in the high accuracy class, indicating that extrapolation mostly affect lower attribute values. A noticeable exception from this trend was observed for GSV predicted using DAP-based model. The mean GSV in the low accuracy class was 19 m^3ha^{-1} greater than the mean GSV in the high accuracy class (i.e. 178.8 versus 159.8 m^3ha^{-1}), showing a possible saturation effect of the model toward the greatest GSV.

The reliability assessment performed through the intersection of precision and accuracy is presented in Table 8 and Fig. 7. More than 86.3% and 83.5% of the predictions are classified as highly reliable for ALS and DAP respectively (Table 8). On the opposite, low reliability accounted for 4.7% (VI, DAP) up to 9.5% (GSV, DAP) of the predictions. Driven by the precision classes, the mean distance increased with decreasing reliability. That said, the low reliability class exhibited mean distances around 2 times greater than the other 2 classes. By comparison with the low precision (Table 6), the low reliability had greater prediction means. For ALS-based model, the GSV in low precision was 21.0 m^3ha^{-1} which increased up to 103.1 m^3ha^{-1} in the low reliability class. Interestingly in DAP, the large GSV found in low accuracy (178.8 m^3ha^{-1}, Table 7) decreased to 93.8 m^3ha^{-1} in the low reliability class (Table 8), indicating that a significant proportion of predictions in this class had low attribute values, and are driven by low precision. Such a result could can also be seen in Fig. 7, which indicated that low precision is predominantly affecting low attribute values while low accuracy is mostly found towards greater attribute values.

5. Discussion

5.1. Model development and assessment

In this study, models based on ALS and DAP included information from the three auxiliary data sources considered. The forest type information obtained from IGN forest map is a key attribute for differentiating structural and functional traits (Irulappa-Pillai-Vijayakumar et al., 2019). Landsat 8 green band acquired between 0.53 and 0.59 μm and short-wave infrared band 1 (1.57–1.65 μm) emerged as the most influential optical remote sensing variables. Astola et al. (2019) found that Landsat 8 visible green band was ranked first for predicting GSV and BA in the boreal forest of Finland. As emphasized by Cohen and Goward (2004), the SWIR1 band provided information about the vegetation and ground humidity and could be associated here with productivity gradients. The higher number of 3D variables selected in the models is associated with the high correlation between the vertical distribution of vegetation and the major forest attributes (Næsset, 2002). ALS data convey a higher amount of information related to the vertical distribution of the vegetation, which could explain why a higher number of ALS metrics (3 metrics) were selected compared to DAP (2 metrics) data (Filippelli et al. 2019). DAP point clouds are restricted to the outer canopy structure and could not provide information about the vertical distribution of vegetation (St-Onge et al., 2008; White et al., 2013). Similar trends in the number of auxiliary variables were also reported by Järnstedt et al. (2012).

For the estimation of forest resources, Chirici et al. (2016) and Cosenza et al. (2021) reported that *k*-NN models with 5 nearest neighbours and Euclidean distance performed better than any other *k*-NN configuration. Here, the selected models were based on the msn distance. The higher performance of msn over Euclidean could be explained by the fact that msn distance integrates information from the field attributes. That said, this could make *k*-NN models less flexible to predict field attributes which were not considered during calibration. From investigating over 260 studies, Chirici et al. (2016), reported an average rRMSE of 37% (±32%) for growing stock volume, biomass and carbon stock. With test errors ranging from 46% (ALS, BA) to 67% (ALS, VI), our results are in the upper range of results reported by Chirichi et al. (2016). The large errors achieved here could be explained by the model selection method, which emphasize on model parsimony (Pestov, 2013), and reduction of overfitting (Valbuena et al. 2017). Indeed, numerous studies included a large number of auxiliary variables in their models (Beaudoin et al., 2014; Nurminen et al., 2013), which tends to reduce the variance at the expense of a possible risk of model overfitting and bias. Part of our error could also be explained by positioning precision of the field plots (Fadili et al., 2019), as well as the complexity of the forest structure studied, with mixture of conifers and hardwood stands and various management strategies from coppice stand to even-aged ones, and uneven-aged stands resulting from the transition from coppice to even-aged stands (Jarret, 2004). Regarding field attributes, the larger errors achieved for VI with respect to both GSV and BA could be explained by the nature of the attribute. VI is computed from volume and the 5 year radial increment, thus including a temporal dimension that is not well captured with the auxiliary data used. As shown in Irulappa-Pillai-Vijayakumar et al. (2019), the estimation of such flux-related attribute could be improved by taking into account auxiliary data from various points in time.

5.2. Statistical inference and mapping

In line with the model's performances, the population estimates from both ALS and DAP models were of the same range. This result is in agreement with previous studies, which also indicated that the performance of DAP and ALS based predictions of forest attributes are similar (Bohlin et al., 2012; Gobakken et al., 2015; Järnstedt et al., 2012; Nurminen et al., 2013). This highlights the capability of DAP to compete

with ALS in the estimation of major forest attributes. The large spatial coverage of aerial surveys and their high renewal rates in many countries makes DAP a serious competitor to ALS when a high resolution DTM is available (St-Onge et al., 2008). The availability of time series of aerial photographs opens up the possibility to further improve the variance of estimates of forest attributes related to fluxes such as VI (Irulappa-Pillai-Vijayakumar et al., 2019) and could also be an integral part of a forest monitoring systems (Ginzler et al., 2019).

Nevertheless, the inflated variance observed for the GSV predicted with ALS was unexpected and questioned the quality of the model at population level. A possible explanation could be found in the *k*-NN distances and amount of extrapolation. The ALS model showed larger mean *k*-NN distances than it's DAP counterpart (Table 8). Also, mean *k*-NN distances are inflating faster with decreasing reliability (i.e.. from high to low) with ALS than with DAP (Table 8). This inflation in distances suggested that the choice of neighbours during the resampling process was more variable with ALS, thus contributing to an increased variance at the population level. This phenomenon was accentuated by the higher proportion of extrapolation with ALS (Table 7). A possible reason for this result may be associated to the nature of the data. As ALS is more sensitive to the inner-crown canopy structure (Filippelli et al., 2019), one might consider that the variability brought in the auxiliary variable space could have introduced uncertainty in the neighbourhood selection during the resampling, contributing to an increase variance at the population level. The results of the inverted models, generated by applying the auxiliary variables selected with one model to the other, and provided in Table A1 and A2, showed that predictions with ALS remained more variables, corroborating this hypothesis.

The high-resolution maps complement the bootstrap model-based estimates, by providing information about the localisation of the forest resources as well as their reliability. These maps can be anticipated to be increasingly relevant for future forest management and planning. The *k*-NN predictions were found to be imprecise in young stands and along forest transitions (Fig. 7), which are not measured in the French NFI. Indeed, with a minimum circumference at breast height of 23.5 cm, forest stands below 10 m height are very poorly described. Extrapolation was also found in those areas. More importantly, with the use of the convex hull approach, we observed that an even higher proportion of the pixels were extrapolated in highly stocked stands. This highlights a possible saturation of the model towards the upper values of the targeted forest attributes. It underlines the need to further study the magnitude of such a bias and the possibilities to correct it. Interestingly, the mean distance to the closest neighbours is significantly different for interpolation and extrapolation, suggesting that the extrapolation distance could serve as a metric to correct extrapolation bias in *k*-NN models (Magnussen et al., 2010).

A limitation of our results comes from the binary assessment of the extrapolation. We used the entire calibration domain to identify the extrapolated pixels, resulting in a binary map. A more robust but time-consuming approach would be to assess extrapolation both at the plot level and population level by taking advantage of the bootstrap properties. Computing extrapolation for each bootstrap sample would provide a continuous evaluation of extrapolation (i.e. varying from 0 = no extrapolation; to the number of bootstrap samples = full extrapolation). Such information could be useful in complement the *k*-NN distances to analyse the relationship between extrapolation and errors at plot level (Meyer and Pebesma 2021), to identify outliers at both plot and population level (Cook, 1977) and develop appropriate bias-correction methods.

6. Conclusion

This research is a step ahead to connect prediction and reliability maps within an inferential framework. The approach relies on bootstrapped predictions to infer population parameters in a model-based framework and provide high-resolution maps of forest attributes with reliability assessment. The approach was developed for *k*-NN models, owing to their capabilities to predict multiple forest attributes in a single model. To account for the incapacity of *k*-NN to extrapolate, we further adapted the convex hull approach to measure the accuracy of prediction. Bootstrapped predictions with precision and accuracy were combined to provide a measure of prediction reliability at the unit level (grid cell or pixel). The high-resolution prediction maps will assist forest managers in the localisation of forest resources, and subsequently, the reliability maps (extrapolation and precision) will allow for judgments about whether the predicted forest attributes map is accurate enough for decision making.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the French PIA project "Lorraine Université d'Excellence", reference ANR-15-IDEX-04-LUE, through the project Impact DeepSurf and the Labex ARBRE (ANR-11-LABX-0002-01). The authors deeply thank David Sterratt and C Bradford Barber for their help with the convex-hull computation with the Quickhull algorithm and the R Geometry package. We also greatly thank Jean-Romain Roussel for his support with R lidR package.

Appendix A. Inverted model outputs

See Tables A1 and A2.

References

Astola, H., Häme, T., Sirro, L., Molinier, M., Kilpi, J., 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.

Barber, C.B., Dobkin, D.P., Huhdanpaa, H., 1996. The quickhull algorithm for convex hulls. ACM Trans. Math. Softw. TOMS 22, 469–483. https://doi.org/10.1145/235815.235821.

Barrett, F., McRoberts, R.E., Tomppo, E., Cienciala, E., Waser, L.T., 2016. A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. Remote Sens. Environ. 174, 279–289. https://doi.org/10.1016/j.rse.2015.08.029.

Beaudoin, A., Bernier, P.Y., Guindon, L., Villemaire, P., Guo, X.J., Stinson, G., Bergeron, T., Magnussen, S., Hall, R.J., 2014. Mapping attributes of Canada's forests at moderate resolution through kNN and MODIS imagery. Can. J. For. Res. 44, 521–532. https://doi.org/10.1139/cjfr-2013-0401.

Bhanu, B., Lin, Y., 2003. Genetic algorithm based feature selection for target detection in SAR images. Image Vis. Comput. 21, 591–608. https://doi.org/10.1016/S0262-8856(03)00057-X.

Blackard, J.A., Finco, M.V., Helmer, E.H., Holden, G.R., Hoppus, M.L., Jacobs, D.M., Lister, A.J., Moisen, G.G., Nelson, M.D., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D.L., Winterberger, K.C., Brandeis, T.J., Czaplewski, R.L., McRoberts, R.E., Patterson, P.L., Tymcio, R.P., 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. Remote Sens. Environ. 112, 1658–1677. https://doi.org/10.1016/j.rse.2007.08.021.

Bohlin, J., Wallerman, J., Fransson, J.E.S., 2012. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. Scand. J. For. Res. 27, 692–699. https://doi.org/10.1080/02827581.2012.686625.

Breidenbach, J., Astrup, R., 2012. Small area estimation of forest attributes in the Norwegian National Forest Inventory. Eur. J. For. Res. 131, 1255–1267. https://doi.org/10.1007/s10342-012-0596-7.

Breidenbach, J., Næsset, E., Gobakken, T., 2012. Improving k-nearest neighbor predictions in forest inventories by combining high and low density airborne laser scanning data. Remote Sens. Environ. 117, 358–365. https://doi.org/10.1016/j.rse.2011.10.010.

Breidenbach, J., McRoberts, R.E., Alberdi, I., Antón-Fernández, C., Tomppo, E., 2021. A century of national forest inventories – informing past, present and future decisions. For. Ecosyst. 8, 36. https://doi.org/10.1186/s40663-021-00315-x.

Brosofske, K.D., Froese, R.E., Falkowski, M.J., Banskota, A., 2014. A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management. For. Sci. 60, 733–756. https://doi.org/10.5849/forsci.12-134.

Castillo-Santiago, M.A., Ricker, M., de Jong, B.H.J., 2010. Estimation of tropical forest structure from SPOT-5 satellite images. Int. J. Remote Sens. 31, 2767–2782. https://doi.org/10.1080/01431160903095460.

Chirici, G., Mura, M., McInerney, D., Py, N., Tomppo, E.O., Waser, L.T., Travaglini, D., McRoberts, R.E., 2016. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. Remote Sens. Environ. 176, 282–294. https://doi.org/10.1016/j.rse.2016.02.001.

Chirici, G., Giannetti, F., McRoberts, R.E., Travaglini, D., Pecchi, M., Maselli, F., Chiesi, M., Corona, P., 2020. Wall-to-wall spatial prediction of growing stock volume based on Italian National Forest Inventory plots and remotely sensed data. Int. J. Appl. Earth Obs. Geoinformation 84, 101959. https://doi.org/10.1016/j.jag.2019.101959.

Coelho, P.S., Pereira, L.N., 2011. A spatial unit level model for small area estimation. Revstat. 9, 155–180.

Cohen, W.B., Goward, S.N., 2004. Landsat's Role in Ecological Applications of Remote Sensing. Bioscience 54, 535–545. https://doi.org/10.1641/0006-3568(2004)054[0535:LRIEAO]2.0.CO;2.

Conn, P.B., Johnson, D.S., Boveng, P.L., Boulinier, T., 2015. On Extrapolating Past the Range of Observed Data When Making Statistical Predictions in Ecology. PLoS ONE 10 (10), e0141416.

Cook, R.D., 1977. Detection of Influential Observation in Linear Regression. Technometrics 19, 15–18. https://doi.org/10.1080/00401706.1977.10489493.

Coops, N.C., Tompalski, P., Goodbody, T.R.H., Queinnec, M., Luther, J.E., Bolton, D.K., White, J.C., Wulder, M.A., van Lier, O.R., Hermosilla, T., 2021. Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends. Remote Sens. Environ. 260, 112477. https://doi.org/10.1016/j.rse.2021.112477.

Cosenza, D.N., Korhonen, L., Maltamo, M., Packalen, P., Strunk, J.L., Næsset, E., Gobakken, T., Soares, P., Tomé, M., 2021. Comparison of linear regression, k-nearest neighbour and random forest methods in airborne laser-scanning-based prediction of growing stock. Forestry 94, 311–323. https://doi.org/10.1093/forestry/cpaa034.

Crookston, N.L., Finley, A.O., 2008. yaImpute: An R Package for kNN Imputation. J. Stat. Softw. 23, 1–16. https://doi.org/10.18637/jss.v023.i10.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.

Durrieu, S., Vega, C., Bouvier, M., Gosselin, F., Renaud, J.-P., Saint-André, L., 2015. Optical remote sensing of tree and stand heights. In: Land Resources Monitoring, Modeling, and Mapping With Remote Sensing, Remote Sensing Handbook. CRC Press, pp. 449–485. https://doi.org/10.1201/b19322.

Esteban, J., McRoberts, R., Fernández-Landa, A., Tomé, J., Næsset, E., 2019. Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. Remote Sens. 11, 1944. https://doi.org/10.3390/rs11161944.

Fadili, M., Renaud, J.-P., Bock, J., Vega, C., 2019. RegisTree: a registration algorithm to enhance forest inventory plot georeferencing. Ann. For. Sci. 76, 30. https://doi.org/10.1007/s13595-019-0814-2.

Filippelli, S.K., Lefsky, M.A., Rocca, M.E., 2019. Comparison and integration of lidar and photogrammetric point clouds for mapping pre-fire forest structure. Remote Sens. Environ. 224, 154–166. https://doi.org/10.1016/j.rse.2019.01.029.

Ginzler, C., Price, B., Bösch, R., Fischer, C., Hobi, M.L., Psomas, A., Rehush, N., Wang, Z., Waser, L.T., 2019. Area-Wide Products. In: Fischer, C., Traub, B. (Eds.), Swiss National Forest Inventory – Methods and Models of the Fourth Assessment, Managing Forest Ecosystems, vol. 35. Springer, Cham, pp. 125–142. https://doi.org/10.1007/978-3-030-19293-8_7.

Gobakken, T., Bollandsås, O.M., Næsset, E., 2015. Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data. Scand. J. For. Res. 30, 73–86. https://doi.org/10.1080/02827581.2014.961954.

Goerndt, M.E., Wilson, B.T., Aguilar, F.X., 2019. Comparison of small area estimation methods applied to biopower feedstock supply in the Northern U.S. region. Biomass Bioenergy 121, 64–77. https://doi.org/10.1016/j.biombioe.2018.12.008.

Gregoire, T.G., Næsset, E., McRoberts, R.E., Ståhl, G., Andersen, H.-E., Gobakken, T., Ene, L., Nelson, R., 2016. Statistical rigor in LiDAR-assisted estimation of aboveground forest biomass. Remote Sens. Environ. 173, 98–108. https://doi.org/10.1016/j.rse.2015.11.012.

Guitet, S., 2018. Analyses du contrôle national : Premières visites. Inventaire Forestier National 153 p.

Habel, K., Grasman, R., Gramacy, R.B., Mozharovskyi, P., Sterratt, D.C., 2019. Geometry: Mesh Generation and Surface Tessellation. R package version 0.4.5. https://CRAN.R-project.org/package=geometry.

Hagolle, O., Huc, M., Villa Pascual, D., Dedieu, G., 2015. A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENμS and Sentinel-2 Images. Remote Sens. 7, 2668–2691. https://doi.org/10.3390/rs70302668.

Hervé, J.-C., Wurpillot, S., Vidal, C., Roman-Amat, B., 2014. L'inventaire des ressources forestières en France : un nouveau regard sur de nouvelles forêts. Rev. For. Fr. 247–260 https://doi.org/10.4267/2042/56055.

Hou, Z., Xu, Q., McRoberts, R.E., Greenberg, J.A., Liu, J., Heiskanen, J., Pitkänen, S., Packalen, P., 2017. Effects of temporally external auxiliary data on model-based inference. Remote Sens. Environ. 198, 150–159. https://doi.org/10.1016/j.rse.2017.06.013.

Irulappa-Pillai-Vijayakumar, D.B., Renaud, J.-P., Morneau, F., McRoberts, R.E., Vega, C., 2019. Increasing Precision for French Forest Inventory Estimates using the k-NN Technique with Optical and Photogrammetric Data and Model-Assisted Estimators. Remote Sens. 11, 991. https://doi.org/10.3390/rs11080991.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. New York: springer. 112, 18. https://doi.org/10.1007/978-1-4614-7138-7.

Järnstedt, J., Pekkarinen, A., Tuominen, S., Ginzler, C., Holopainen, M., Viitala, R., 2012. Forest variable estimation using a high-resolution digital surface model. ISPRS J. Photogramm. Remote Sens. 74, 78–84. https://doi.org/10.1016/j.isprsjprs.2012.08.006.

Jarret, P., 2004. Guide des Sylvicultures: Chênaie Atlantique; Office National des Forêts: Paris, France.

Kangas, A., Astrup, R., Breidenbach, J., Fridman, J., Gobakken, T., Korhonen, K.T., Maltamo, M., Nilsson, M., Nord-Larsen, T., Næsset, E., Olsson, H., 2018. Remote sensing and forest inventories in Nordic countries – roadmap for the future. Scand. J. For. Res. 33, 397–412. https://doi.org/10.1080/02827581.2017.1416666.

Longford, N.T., 2005. Editorial: Model selection and efficiency-is "Which model?" the right question? J. R. Statist. Soc. Ser. A 168, 469–472. https://doi.org/10.1111/j.1467-985X.2005.00366.x.

Magnussen, S., McRoberts, R.E., Tomppo, E.O., 2009. Model-based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. Remote Sens. Environ. 113, 476–488. https://doi.org/10.1016/j.rse.2008.04.018.

Magnussen, S., Tomppo, E., McRoberts, R.E., 2010. A model-assisted k-nearest neighbour approach to remove extrapolation bias. Scand. J. For. Res. 25, 174–184. https://doi.org/10.1080/02827581003667348.

Magnussen, S., McRoberts, R.E., Breidenbach, J., Nord-Larsen, T., Ståhl, G., Fehrmann, L., Schnell, S., 2020. Comparison of estimators of variance for forest inventories with systematic sampling - results from artificial populations. For. Ecosyst. 7, 17. https://doi.org/10.1186/s40663-020-00223-6.

Mascaro, J., Detto, M., Asner, G.P., Muller-Landau, H.C., 2011. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. Remote Sens. Environ. 115, 3770–3774. https://doi.org/10.1016/j.rse.2011.07.019.

McRoberts, R.E., 2012. Estimating forest attribute parameters for small areas using nearest neighbors techniques. For. Ecol. Manag. 272, 3–12. https://doi.org/10.1016/j.foreco.2011.06.039.

McRoberts, R.E., Magnussen, S., Tomppo, E.O., Chirici, G., 2011. Parametric, bootstrap, and jackknife variance estimators for the k-Nearest Neighbors technique with illustrations using forest inventory and satellite image data. Remote Sens. Environ. 115, 3165–3174. https://doi.org/10.1016/j.rse.2011.07.002.

McRoberts, R.E., Næsset, E., Gobakken, T., 2015. Optimizing the k-Nearest Neighbors technique for estimating forest aboveground biomass using airborne laser scanning data. Remote Sens. Environ. 168, 252–264. https://doi.org/10.1016/j.rse.2015.02.026.

McRoberts, R.E., Næsset, E., Saatchi, S., Quegan, S., 2022. Statistically rigorous, model-based inferences from maps. Remote Sens. Environ. 279, 113028. https://doi.org/10.1016/j.rse.2022.113028.

Meyer, H., Pebesma, E., 2021. Predicting into unknown space? Estimating the area of applicability of spatial prediction models. Methods. Ecol. Evol. 12, 1620–1633. https://doi.org/10.1111/2041-210X.13650.

Moeur, M., Stage, A.R., 1995. Most similar neighbor: An improved sampling inference procedure for natural resource planning. Forest Science 41, 337–359.

Moser, P., Vibrans, A.C., McRoberts, R.E., Næsset, E., Gobakken, T., Chirici, G., Mura, M., Marchetti, M., 2017. Methods for variable selection in LiDAR-assisted forest inventories. Forestry 90, 112–124. https://doi.org/10.1093/forestry/cpw041.

Næsset, E., 2002. Determination of Mean Tree Height of Forest Stands by Digital Photogrammetry. Scand. J. For. Res. 17, 446–459. https://doi.org/10.1080/02827580232045469.

Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., Larsson, S., Nilsson, L., Eriksson, J., Olsson, H., 2017. A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory. Remote Sens. Environ. 194, 447–454. https://doi.org/10.1016/j.rse.2016.10.022.

Nurminen, K., Karjalainen, M., Yu, X., Hyyppä, J., Honkavaara, E., 2013. Performance of dense digital surface models based on image matching in the estimation of plot-level forest variables. ISPRS J. Photogramm. Remote Sens. 83, 104–115. https://doi.org/10.1016/j.isprsjprs.2013.06.005.

Pestov, V., 2013. Is the k -NN classifier in high dimensions affected by the curse of dimensionality? Comput. Math. Appl. 65, 1427–1437. https://doi.org/10.1016/j.camwa.2012.09.011.

Puliti, S., Breidenbach, J., Schumacher, J., Hauglin, M., Klingenberg, T.F., Astrup, R., 2021. Above-ground biomass change estimation using national forest inventory data with Sentinel-2 and Landsat. Remote Sens. Environ. 265, 112644. https://doi.org/10.1016/j.rse.2021.112644.

Rahlf, J., Breidenbach, J., Solberg, S., Næsset, E., Astrup, R., 2017. Digital aerial photogrammetry can efficiently support large-area forest inventories in Norway. Forestry 90, 710–718. https://doi.org/10.1093/forestry/cpx027.

Rahlf, J., Hauglin, M., Astrup, R., Breidenbach, J., 2021. Timber volume estimation based on airborne laser scanning — comparing the use of national forest inventory and forest management inventory data. An. For. Sci. 78, 49. https://doi.org/10.1007/s13595-021-01061-4.

Robert, N., Vidal, C., Colin, A., Hervé, J.-C., Hamza, N., Cluzeau, C., 2009. Developement of France's national forest inventory. In: Tomppo, E., Gschwantner, T., Lawrence, M., McRoberts, R.E. (Eds.), National forest Inventories – Pathways for Common Reporting. Springer, Cham, pp. 207–221.

Roussel, J.-R., Auty, D., Coops, N.C., Tompalski, P., Goodbody, T.R.H., Meador, A.S., Bourdon, J.-F., de Boissieu, F., Achim, A., 2020. lidR: An R package for analysis of

Airborne Laser Scanning (ALS) data. Remote Sens. Environ. 251, 112061. https://doi.org/10.1016/j.rse.2020.112061.

Rupnik, E., Daakir, M., Pierrot Deseilligny, M., 2017. MicMac – a free, open-source solution for photogrammetry. Open Geospatial Data Softw. Stand. 2, 1–9. https://doi.org/10.1186/s40965-017-0027-2.

Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyyppä, J., 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. Remote Sens. Environ. 158, 431–440. https://doi.org/10.1016/j.rse.2014.11.020.

Saarela, S., Wästlund, A., Holmström, E., Mensah, A.A., Holm, S., Nilsson, M., Fridman, J., Ståhl, G., 2020. Mapping aboveground biomass and its prediction uncertainty using LiDAR and field data, accounting for tree-level allometric and LiDAR model errors. For. Ecosyst. 7, 43. https://doi.org/10.1186/s40663-020-00245-0.

Souchon, J.-P., Thom, C., Meynard, C., Martin, O., Pierrot-Deseilligny, M., 2010. The IGN CAMv2 System. Photogramm. Rec. 25, 402–421. https://doi.org/10.1111/j.1477-9730.2010.00601.x.

Ståhl, G., Saarela, S., Schnell, S., Holm, S., Breidenbach, J., Healey, S.P., Patterson, P.L., Magnussen, S., Næsset, E., McRoberts, R.E., Gregoire, T.G., 2016. Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. For. Ecosyst. 3, 5. https://doi.org/10.1186/s40663-016-0064-9.

St-Onge, B., Vega, C., Fournier, R.A., Hu, Y., 2008. Mapping canopy height using a combination of digital stereo-photogrammetry and lidar. Int. J. Remote Sens. 29, 3343–3364. https://doi.org/10.1080/01431160701469040.

Tewari, V.P., Kleinn, C., 2015. Considerations on capacity building for national forest assessments in developing countries – with a case study of India. Int. For. Rev. 17, 244–254. https://doi.org/10.1505/146554815815500633.

Tomppo, E., Haakana, M., Katila, M., Peräsaari, J., 2008. Multi-source national forest inventory: methods and applications. In: Managing Forest Ecosystems, vol. 18. Springer, p. 374.

Valbuena, R., Hernando, A., Manzanera, J.A., Görgens, E.B., Almeida, D.R.A., Mauro, F., García-Abril, A., Coomes, D.A., 2017. Enhancing of accuracy assessment for forest above-ground biomass estimates obtained from remote sensing via hypothesis testing and overfitting evaluation. Ecol. Modell. 366, 15–26. https://doi.org/10.1016/j.ecolmodel.2017.10.009.

Vega, C., Renaud, J.-P., Sagar, A., Bouriaud, O., 2021. A new small area estimation algorithm to balance between statistical precision and scale. Int. J. Appl. Earth Obs. Geoinformation 97, 102303. https://doi.org/10.1016/j.jag.2021.102303.

White, J., Wulder, M., Vastaranta, M., Coops, N., Pitt, D., Woods, M., 2013. The Utility of Image-Based Point Clouds for Forest Inventory: A Comparison with Airborne Laser Scanning. Forests 4, 518–536. https://doi.org/10.3390/f4030518.

Zhao, P., Lu, D., Wang, G., Wu, C., Huang, Y., Yu, S., 2016. Examining Spectral Reflectance Saturation in Landsat Imagery and Corresponding Solutions to Improve Forest Aboveground Biomass Estimation. Remote Sens. 8, 469. https://doi.org/10.3390/rs8060469.

Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sens. Environ. 128, 289–298. https://doi.org/10.1016/j.rse.2012.10.017.