Big Data bases and applications

Mapping and SWOT analysis of existing and future Big Data sources

28 February 2017

Version	Date	Written by	Reviewed by	Approved by
1.0	28 February 2017	Erik Willén, Sima Mohtashami, Aron Davidsson, Gustav Friberg (Skogforsk), Harri Lindeman, Kari Väätäinen, Matti Sirén (Luke), Jarmo Hämäläinen, Tapio Räsänen (Metsäteho), Morgan Vuillermoz, Christophe Ginet (FCBA)	Tomas Nordfjell (SLU)	Jori Uusitalo (Luke)
1.1	5 April 2017			

Table of contents

1. EFFORTE project objectives

EFFORTE is a research and innovation project providing the European forestry sector with new knowledge and knowhow that will significantly improve the possibilities of forest enterprises to assemble and adopt novel technologies and procedures.

The project aims at enhancing the efficiency of silviculture and harvesting operations; increasing wood mobilization and annual forest growth; increasing forest operations' output while minimizing environmental impacts; and reducing fuel consumption in the forest harvesting process by at least 15%.

The project is based on three key elements of technology and knowhow:

1) Basic understanding of fundamentals of **soil mechanics and terrain trafficability** is a crucial starting point to avoid soil disturbances, accelerate machine mobility and assess persistence of soil compaction and rutting. The key findings and recommendations of trafficability related to EFFORTE can immediately be adapted in all European countries.

2) Due to decreasing Cost-competitiveness of manual work and maturity of technology it is now perfect time to realize the potential of **mechanization in silvicultural operations**. EFFORTE pursues for higher productivity and efficiency in silvicultural operations such as tree planting and young stand cleaning operations.

3) 'Big Data' (geospatial as well as data from forestry processes and common information e.g. weather data) provides a huge opportunity to increase the efficiency of forest operations. In addition it adds new possibilities to connect knowledge of basic conditions (e.g. trafficabillity), efficient silviculture and harvesting actions with demand and expectations from forest industries and the society. Accurate spatial information makes it possible for forestry to move from classic stand-wise management to precision forestry, i.e. micro stand level, grid cell level or tree-by-tree management. EFFORTE aims at achieving substantial influence to the **implementation and improved use of Big Data within Forestry** and through this increase Cost-efficiency and boost new business opportunities to small and medium size enterprises (SME) in the bioeconomy.

EFFORTE researchers will develop and pilot precision forestry applications that, according to the industrial project partners, show the greatest potential for getting implemented immediately after the project.

2. Mapping and SWOT analysis of existing and future Big Data sources

This report includes:

- Existing data sources (country wise) of relevance for planning of operations and wood supply in the forest industry. Everything is identified and described with regards to current and future accessibility, data quality, level of standardisation, and existing and future fields of application.

-Potential future data sources (country wise) are identified and assessed with regards to feasibility.

- A SWOT analysis (country wise) of most relevant existing and potential future data sources.

- Conclusions including the outlook and the European dimension based on the experiences in Sweden, Finland and France.

The following categories of data sources are included in the analysis:

Category	Example of data sources	Specific features
Remote sensing	Earth Observation data, aerial imagery, lidar data, UAV	Raster data or point clouds
Elevation and surface models	Digital elevation models, Digital surface models	Raster data
Map data	Topographic maps, soil maps, cadastral data, wet area maps	Raster or vector data, complete coverage
Thematic databases	Road databases, Protected areas, Natura2000-sites, cultural heritages	Vector data, partial coverage
In-situ data	Field measurements, forest stand databases, harvester and forwarder data	Fully or partially collected in field
Meteorological data	Weather data, forecasts	
Forest estimates	Estimations of forest parameters such as basal area, mean diameter or forest volume, based on remote sensing	Raster data
Statistics	Transportation statistics, industry demand, forest statistics	Restricted access for some statistics

All existing data sources are listed and described with the following parameters. The future data sources are described with the same parameters except for "Existing applications in planning and wood supply" that are assumed to be none for future data sources. Certain parameters in the tables are not known or have not been found and are then "non-applicable (n.a.)".

Parameter	Description
Data source	Name of data source
Producer	Producer of data
Coverage	Spatial coverage, e.g. full coverage of Sweden or covering area of forest owner
Update frequency	For example yearly or every 10th year or not planned
Delivery format	Raster/vector/WMS/WFS. Spatial resolution in meter, if applicable
Accessibility	Free/cost/restricted
Data quality	Comments on data quality, e.g. accuracy
Level of standardization	High, medium or low
Existing applications in planning and wood supply	List existing application areas
Potential applications in planning and wood supply	List potential application areas
Reference	Link to webpage or similar

3. Existing Big Data sources

3.1. Overview

Geospatial as well as data from forestry processes and common information e.g. weather data provides a huge opportunity to increase the efficiency of forest operations. A general trend is increased availability of different kind of data sources both in terms of the amount of data and reduced costs.

The INSPIRE (Infrastructure for Spatial Information in Europe) directive[1] requires an infrastructure for geodata that allows users to search for and to visualize public data without cost in all member states. There might be cost for downloading data, but in many European countries public geodata is also downloaded for free. Public investments get available for all users. This is the case in e.g. Netherlands, Denmark and Finland while in some countries, as Sweden, users still need to pay for downloading public geodata. However, it is most likely that public geodata will be free of charge in these countries within the near future. Research organizations most often get data free of charge for R&D in all countries.

There are however some geospatial data not part of the INSPIRE directive and connected to either charges or confidentiality. Refined decision support provided by consultants may in many cases be innovative and customer-oriented, but comes with a fee. Many of the available derived datasets produced from combinations of big data sources are offered by consultants and this also drives the innovation in a positive way.

Other datasets, related to forest operations or some statistics, are confidential. Harvester data belongs to the forest owner or entrepreneur and detailed statistics about wood supply is seldom public. Soon, even more data from machines in forest operations and transportation may be utilized, but this is also most likely to be restricted in terms of availability. This is not necessarily a “show stopper” for the development of refined decision support, but will include data owners in the R&D efforts.

The focus on big datasets listed are linked to the objectives of the EFFORTE project, i.e. support to terrain trafficability, silvicultural and harvesting forest operations. The current state of the art varies in each country this is partly reflected in each country table. The access to detailed data is also a prerequisite for the decision support to be developed in the EFFORTE project and this access vary along the countries which may explain the different inputs.

[1] http://inspire.ec.europa.eu/

3.2. Sweden

Introduction

The access to Big Databases in Sweden is generally very good with lots of public data available both for download or through WMS- services. The Swedish response to the INSPIRE directive is the website www.geodata.se where it is possible to search, visualize and to download data from a wide range of agencies. Of special interest are Lantmateriet (the Swedish mapping, cadastre and land registration authority), Swedish Forest Agency and SMHI (Met office). However, most high-resolution maps and images are not free for download. Orthophotos, lidar data and detailed map data from Lantmateriet do have fees now. There is an agreement with all research organisations to use data for R&D purposes free of charge, but fees are a limiting factor especially for SMEs.

Sweden also holds the Saccess database (https://saccess.lantmateriet.se) where nationwide ortho-corrected cloud-free satellite data from 1970th to recent years are available for download without costs. This will during 2017-2018 be updated to "Svea" database acting as a ESA Sentinel collaborative ground segment.

For map data and thematic databases new products constantly enters the market. The focus is products with large area coverage and to meet the EFFORTE objectives.

Remote sensing

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardi-zation	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Orthophoto	Lantmateriet	Nationwide	2-6 years	Raster, 0,25-0,5 m spatial resolution; pan, colour, colour-infrared	Cost	Root mean square error (RMS) in location 0,3 m	High	Detailed planning in production and considerations, damage assessment, monitoring	Detailed planning in production and considerations, damage assessment, monitoring	www.lantmateriet.se

Orthophoto	Service providers	partly	On request	Raster, 0,05-0,25 m spatial resolution	Cost	Depending on ground control points, similar as nationwide orthophoto	High	Detailed planning in production and considerations, damage assessment, monitoring	Detailed planning in production and considerations, damage assessment, monitoring	e.g: www.blomasa.com
Historic Orthophoto (Year 1930-1990)	Lantmateriet	Nationwide	n.a.	Raster, 0,5 m	Cost	RMS location error 2 m	Medium		Use of time stamps to map forest conditions, Old forest mapping	www.lantmateriet.se
Very High resolution satellite data	WorldView, Pleijades, Spot 6/7, GeoEye, (RapidEye)	Partly	On request	Raster, 0,3-2 (6) m	Cost	RMS location error 0,5 m	High	Detailed planning in production and considerations, damage assessment, monitoring	Detailed planning in production and considerations, damage assessment, monitoring, Forest estimates	e.g: www.metria.se www.planet.com
High resolution optical satellite data	ESA (Sentinel 2), USGS (Landsat 8), Saccess	Nationwide	1-8 coverage yearly	Raster, 10-30 m	Free	About 0,5 pixel (5-15 m) for orthocorrected products	High	Mapping of forest types and changes	Monthly monitoring during vegetation period for e.g insect damage	e.g: Sentinel 2[2] Landsat [3] www.lantmateriet.se

[2] https://sentinels.copernicus.eu/web/sentinel/home
[3] https://landsat.usgs.gov/

	database									
Very High resolution radar satellite data	e.g Astrium (TerraSAR), Cosmo-skymed	Nationwide	1-12 coverage yearly	Raster, 0,25-4 m	Cost	0,5-1 pixelsize	High	Mapping of forest types and changes	Mapping of forest types and changes, digital surface model	e.g. Terra SAR-X [4] Cosmo Skymed [5]
High resolution radar satellite data	ESA (Sentinel 1)	Nationwide	1-12 coverage yearly	Raster, about 10-20 m	Free		High	Possible mapping of forest and changes	Possible change monitoring	Sentinel 1[6]
Airborne lidar data	Lantmateriet	Nationwide	1 coverage	0,5-1 points/m^2	Cost		High	Digital terrain model, forest estimates	Forest estimates, mapping of environmental and cultural considerations	www.lantmateriet.se
Airborne lidar data	Service providers	Partly	On request	0,5-15 points/m^2	Cost		High	Digital terrain model, forest estimates	Digital terrain model, forest estimates (single tree), detailed mapping of environmen	e.g www.terratec.se / www.cowi.se

[4] http://www.intelligence-airbusds.com/terrasar-x/
[5] http://www.e-geos.it/cosmo-skymed.html/
[6] https://sentinels.copernicus.eu/web/sentinel/home

									tal and cultural consideratio ns	
Unmanned aerial vehicle, (UAV)	Service provider s	Partly	On request	0,05-0,5 m	Cost	Depending on ground truth	Medium	Detailed planning in production and consideratio ns, damage assessment, monitoring	Detailed planning in production and consideratio ns, damage assessment, monitoring, forest estimates	e.g www.swescan.se www.metria.se
Terrestrial laser scanning	Service provider s	Sample plots – partly coverage	On request	About 1 000 000 points/seco nd	Cost		Low	R&D	Detailed mapping of stems and local elevation models, in-situ reference in forest estimates	e.g. www.terratec.se
Airborne hyperspectr al scanner, including gamma radiation	Service provider s	Partly	On request	Raster, 0,2-2 m	Cost		Low	R&D	Detailed mapping of soil, forest and vegetation	e.g. www.terratec.se

Elevation and surface models

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Surface models from orthophoto (2,5D modell), without colors	Lantmateriet	Nationwide	1/3 of the country, gets updated Yearly	LAZ 1.2 (point data format 2)	Cost	0,5-1 m resolution, Average error in Z-direction is expected to be ca 0,4-0,8 m	High		Visualization of the landscape and identification of changes, growth of forest,	http://www.lantmateriet.se/
Surface models from IRF orthophoto, with colors	Lantmateriet	Nationwide	1/3 of the country, gets updated Yearly	LAZ 1.2 (point data format 2)	Cost	0,5-1 m resolution, Average error in Z-direction is expected to be ca 0,4-0,8 m	High		Visualization of the landscape and identification of changes, growth of forest,	http://www.lantmateriet.se/

GRID 2+	Lantmateriet	Nationwide	n.a.	Raster layers ASCII grids, point clouds, WMS	Cost	2m resolution raster, Precision: 0,1 m in Z and 0,3 m in X-Y, worse in hilly areas	High	Planning of roads, estimation of hydrological situation of the terrain,	Further optimization of road planning, raw material exchange, environmental conservations areas	http://www.lantmateriet.se/
GRID 50+	Lantmateriet	Nationwide	n.a.	GeoTiff	Free	50m resolution raster with 1m precision in Z direction, worse in hilly areas	High	For geographic correction of satellite images, make gravity point determinations, make general estimations, route planning		http://www.lantmateriet.se/

Global DEM (ASTER)	The Ministry of Economy, Trade, and Industry (METI) of Japan, NASA (USA)	Worldwide	First version was provided 2009, the second one 2011	TIFF, PNG	Free	30-meter postings and 1 x 1 degree tiles	Medium		Cross-border forest analyses	ASTER[7]
Surface model UAV	Surface providers	On request	n.a.	Point cloud, DSM, 3D, orthofoto	Cost	About 3 cm x, y and about 4 cm i Z RMS	Medium		For detection of surface change, survey studies, etc	For example: Swescan[8]

[7] https://asterweb.jpl.nasa.gov/gdem.asp
[8] http://www.swescan.se/tjanster/flygmatning-inmatning/

Map data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Depth to water maps	Service providers, Swedish Forest Agency	Nationwide	n.a	Raster, 2 m, (WMS)	Cost	Varied accuracy depending on weather and type of soil	Medium	Harvest planning. Avoid rutting on soil and water. Use during forest operations	Be used as risk-index for rutting on soil and water, decision support for forest planting	www.skogsstyrelsen.se www.cowi.se www.metria.se www.foran.se
Erosion map (risk areas for erosion)	SGI	Nationwide	On request	Raster, WMS	Free	n.a.	High		Consideration during planning of forest operations	www.swedgeo.se
Soil maps	SGU	Nationwide	Continously (slowly)	WMS Databases	Free	Varied resolution over Sweden	High	Harvest planning. Avoid rutting on soil and water. Harvesting operations	Decision support for forest planting, road planning, avoid rutting	www.sgu.se
Swedish land cover	Swedish	Nationwi	No	Raster,	Free	25 m	High	Calculate	Mapping of	Swedish

data	Environmental Protection Agency	de		WMS		resolution (71% thematic accuracy)		Nationwide forest coverage	environmental consideration	Environmental Protection Agency[9]
Topographic map (cadastral maps, roads, streams and water-courses)	Lantmateriet	Nationwide	Varying between types of information	Raster, Vector, Map scale about 1:10 000	Cost	Maximum 2 m	Medium	Maps for orientation, road navigation, cadastral maps etc.	Maps for orientation, road navigation, cadastral maps etc.	www.lantmateriet.se
Topographic map (Medium resolution)	Lantmateriet	Nationwide	Yearly	Raster WMS, map scale about 1:50 000 and less	Free	-	Medium	Maps for orientation		www.Lantmateriet.se

[9] www.naturvardsverket.se

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
FMIS registration (contains position and description of cultural heritages)	Swedish Nation-wide Heritage Board (RAÄ)	Nation-wide	Continuous	Vector, database	Free	Varying quality from region to region	High	Protecting known cultural heritages under planning operations	Improvement in detecting undiscovered heritages in forest	http://www.raa.se/
Protected areas: Nationwide parks, nature reserves, Natura2000	Swedish Environ-mental Protection Agency (SEPA)	Nation-wide	Continuous	Map in form of viewing services	Free	Varying quality from region to region	High	Identification of areas with high environmental values		SEPA[10]
Protected	Swedish	Nation-	Continuo	Vector,	Free	Polygons	High	Identificati		www.skogsstyrelsen.

[10] http://www.naturvardsverket.se

areas: key habitats, swamp forest, list of species within the forest, contracted nature reserves	forest Agency, SKS	wide	us	WMS		, points and lines of varied precision		on of protected areas		se
Nature types and biotopes	Swedish Environ-mental Protection Agency	Nation-wide	Continuous	Vector, WMS	Free	Polygons, points and lines of varied precision	High	Identification of protected species		SEPA[11]
Harvested areas	Swedish forest Agency, SKS	Nation-wide	Yearly	Vector, WMS	Free		High	Yearly maps of forest cuttings		www.skogsstyrelsen.se
Harvesting notifications	Swedish forest Agency, SKS	Nation-wide	Continuous	Vector, WMS	Free		High	Cutting notifications		www.skogsstyrelsen.se
NVDB Nationwide road database	Swedish Transport Administration	Nation-wide	Daily	Vector Databases, WMS,	Free	Varying quality, pulic roads better than private roads	High	NVDB-data are used in many ways, for example in GPS navigators, route-planning		http://www.nvdb.se/en

[11] http://www.naturvardsverket.se/

BaTMan (Bridges, tunnels and other infrastructure connected to roads)	Swedish Transport Administration	Nation-wide	Continuous	Database	Free, Regrestation needed	Varying quality, some information is missing	Medium	Management systems for bridges, tunnels, ferry piers and support structures for roads and railways. works required a user.		BaTMan[12]
Riparian zones (land cover maps along water bodies and rivers)	European Environment Agency, EEA	Pan-European	3-5 years	Raster, minimum mapping unit 0,5 hectare	Free	Thematic accuracy 89%	Medium		Planning of environ-mental considerations	http://land.copernicus.eu/
Green linear elements	European Environment Agency, EEA	Pan-European	3-5 years	Raster, linear elements minimum 100 m length	Free	Thematic accuracy 98 %	Medium		Planning of environ-mental considerations	http://land.copernicus.eu/
Green infrastructure (database with core	Swedish Environ-mental Protection Agency	Nation-wide	No	Vector	Free	n.a. (based on a mix of best available	Low		Planning of environ-mental considerations	www.naturvardsverket.se

[12] https://batman.trafikverket.se/externportal

environmental values)	(SEPA)					sources)				
Mineral resources	The Geological Survey of Sweden	Nation-wide	Continuous	Vector	Free	n.a. (based on a mix of best available sources)	Medium		Planning on road construction	www.sgu.se
Forest fire maps: risk of ignition and soil moisture (HBV), contagion and fire behaviour (FWI)	SMHI (Met office)	Nation-wide	Daily	11 km cells	Free	n.a. (based on a mix of best available sources)	High		Decision support prior to forest operations	http://www.smhi.se/
High resolution layer – Forest	European Environment Agency, EEA	Pan-European	3-5 years	20 m resolution	Free		Medium		Mapping of environ-mental considerations	http://land.copernicus.eu/
Thinning need (calculation to find need for thinning)	Swedish forest Agency	Nation-wide	No	WMS, 12,5 m raster	Free viewing	n.a (based on a mix of best available sources)	Low	Find need for thinning		www.skogsstyrelsen.se
Deciduous forest mapping	Service Provider	Nation-wide	On request	Raster	Cost	4 classes, > 75 % accuracy	Medium	Forest certification		www.metria.se

In-situ data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Harvester data	Each machine or SDC	Clear-cut-area (machine) Regional to Nationwide level (SDC)	Daily per machine	.hpr (.xml) .pri (ascii)	Restricted	High Medium	High (Standford2010) Medium	SilviA, Forest company's specific logistics system	Forest estimation of neighborhood stands, real-time monitoring	www.sdc.se
Forwarder data	Each machine or SDC	Clear-cut-area (machine) Regional to Nationwide level (SDC)	Daily per machine	.fpr (.xml) .pri (ascii)	Restricted	High Medium	High (Standford2010) Medium	SilviA, SilviA, Forest company's specific logistic system	real-time monitoring	www.sdc.se
Forest stand databases	Forest-owning companies	Regional, covering area of forest owner	Continuously	Vector	Restricted	Quality is related to last visiting time to a single forest stand in the	Medium (high)	Forest company's specific planning system		Forest companies

						database.				
Field measurements	Forest-owning companies, Wood-purchasing companies	sub-regional	Continuously	Vector	Cost, Restricted	Quality is related to type of inventory method	Medium to high	Forest company's specific planning system		Forest companies
LUCAS field survey	Eurostat	European	3 year interval		Free		Low		Reference data in forest estimates	http://ec.europa.eu/eurostat/
ÖSI (forest management plans)	Swedish Forest Agency	Nationwide	Old inventory, data collected 1983-1992	Scanned old maps, raster	Free	Low	Low	Reference for timber purchase		www.skogsstyrelsen.se
Nationwide forest Inventory	SLU (Swedish University of Agricultural Sciences)	Nationwide	Yearly	Vector (plots)	Free, Restricted	High quality measurements	Medium	Field reference		www.slu.se

Meteorological data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Precipitation (daily forecast & mean)	SMHI	Nationwide	Daily to yearly	Vector	Free	High	High	Breeding of plants	Avoid rutting	www.smhi.se
Temperature (daily forecast & mean)	SMHI	Nationwide	Daily to yearly	Vector	Free	High	High	Breeding of plants		www.smhi.se
Last spring frost	SMHI	Nationwide	-	Vector	Free	High	High	Breeding of plants	Planning of forest planting	www.smhi.se
Climate predictions (A few climate variables)	Worldclim	World	-	Raster	Free	Resolution: 10 minutes - 30 seconds	Medum	Breeding of plants	Predict forest growt.	www.worldclim.org
Frozen ground	Swedish Transport Administration	Nationwide (separate measuring stations)	Daily - Yearly	Vector, Web map	Free	High	High	Logistics planing & Planing of harvestoperations.		www.trafikverket.se
Wind (daily forecast & meanvalues)	SMHI	Nationwide	Daily - yearly	Vector Databas Web	Free	Interpolated values	High	Forest operations		www.smhi.se

Forest estimates

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Skogliga grunddata (lidar derived forest estimates such as basal area, mean diameter and volume)	Swedish forest Agency	Nation wide	1 coverage (no planned update)	12,5 m raster	Free	< 20 % Mean average error in estimates	Medium	Operational planning, timber purchase support, environmental considerations	Operational planning, timber purchase support environmental considerations	www.skogsstyrelsen.se
SLU forest map with forest estimates based on satellite data and NFI	Swedish university of agricultural sciences	Nation wide	5- year interval	25 m raster	Free	Low on pixel level, High > 100 hectares	Medium	Strategic or tactical planning.	Strategic or tactical planning.	www.slu.se
Tree cover density	EEA	Pan-European	3-5 year interval	20 m raster	Free	> 90 % accuracy	Medium		Strategic or tactical planning.	http://land.copernicus.eu/
Tree height maps (derived from lidar)	Swedish Forest Agency/ Service Providers	Nation wide	1 coverage	2 m, raster	Cost (Free for research)	High	Medium	Operational planning	Operational planning, mapping of environmental considerations	www.skogsstyrelsen.se

Statistics

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Price-calculated transports	SDC	90% of Swedish wood-supply transports	Daily	ASCII	Restricted	High	Medium	monitoring of forest transportation, wood-flows maps		www.sdc.se
NFI , The Swedish Nationwide Forest Inventory	Swedish university of agricultural sciences	Nationwide	Yearly	WMS or vector	Free and restricted	High	High	Forest statistics		SLU[13]
Wood consumption statistics	SDC	Nationwide All pulp and board industries as well as approximately 98% of sawmills with production	Yearly	ASCII	Free	High	High	Monitoring Wood consumption		www.sdc.se

[13] www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-Nationwide-forest-inventory/

		greater than 1000 m^3								
Forest statistics	Swedish Forest Agency	Sweden, statistics from Swedish Forest Agency together with other statistic producers.	Yearly	PDF, Excel, Text	Free	High	High	Forest statistics		Swedish Forest Agency[14]
Cost-index forest operations	Skogforsk	Nationwide	Yearly	PDF	Free	Good	High		Strategic or tactical planning, Cost averages	Skogforsk

[14] http://www.skogsstyrelsen.se/Myndigheten/Statistik/Skogsstatistisk-Arsbok/

3.3. Finland

Introduction

Finland is a society strongly based on information and its utilization. Data systems and the information they contain are integral to the operation of the administration, and the administrative services have been extensively automated and made available in electronic format. Information between organisations is increasingly transmitted digitally, and the Nationwide basic registers are also widely utilized, for example. New information or an automated process usually opens new views for the use of information and developing operations. (http://vm.fi/en)

New legislation for opening big data which has been collected by public funding has opened wide possibilities for novel big data applications. For forestry sector, various organizations have opened their databases for public open data. Data is defined as open data, if data is legally, technically and freely available. As an example, Nationwide land survey provides extensive open data sources, for example mapping including airborne laser scanning (ALS) data. In the following Table 1, several open data providers are listed.

Table 1. Open data providers, links and abbreviations.

Data provider	Abbreviation	Reference
Finnish Environmental Institute	SYKE	www.syke.fi/en-US
Finnish Meteorological Institute	FMI	http://en.ilmatieteenlaitos.fi/
Geologigal Survey of Finland	GTK	www.en.gtk.fi
Natural Resources Institute Finland	Luke	www.luke.fi/en/
Metsähallitus, State Forest Enterprice	MH	www.metsa.fi/web/en
Nationwide Land Survey of Finland	NLS	www.maanmittauslaitos.fi/en
Suomen Metsäkeskus, Finnish Forest Centre	MK	www.metsakeskus.fi/en

Various projects have been launched to enhance the utilization of forest Big data in forestry. Recently ended TEKES-funded "Data to Intelligence" programme included "Forest Big Data"-project, which was the first big effort in this theme (Hämäläinen 2016[15]). Government of Finland has selected "Wood on the move and new products from the forests" as a spear head project including to the action plan, where the initial aim is to boost the utilization of forest Big Data and enhance application development around the Big Data. In addition, Forest industries compiled the Vision for the efficient wood supply 2025, where digitalization and industrial internet are expected to offer significant tools for achieving the targets (Rajala et al. 2015[16]).

[15] Hämäläinen, J. 2016. (editor). Kohti puuhuollon digitaalisaatiota. Forest Big Data –hankkeen päätuloksia. Metsätehon tuloskalvosarja 11/2016

[16] Rajala, P, T., Kääriäinen, H., Laitinen, O., Niemelä, T., Pajuoja, H., Väkevä, J. & Hämäläinen, J. 2015. Vision: Efficient Wood Supply 2025. 1–33 slides. Available: http://www.metsateho.fi/wp-content/uploads/Tehokas-puuhuolto_2025_EN_Efficient_Wood_Supply.pdf

Remote sensing

Data source	Producer	Cover age	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Airborne laser scanning (ALS)	NLS, MH, SMK, Forest companies, ALS operators etc.	Partial	NLS no updates planned (NLS)	LAZ, Vector	Open data (NLS)/ Cost/ Restricted	Varying	High	Forest estimates, elevation models, slope models, wetness index	Estimates of trafficability, operator tutoring, operations planning	NLS[17]
Orthophoto	NLS, MH, SMK, Forest companies, Aerial photo-graphy operators etc.	Nation wide	Every 3 to 10 years (NLS)	Raster	Open data (NLS)/ Cost/ Restricted	Varying	High	Visual forest management and operations planning, Forest estimates,	Digital operations planning	e.g: NLS[18]
Satellite data (see 5.2)										
Aerogeophysical surveying (e.g. gamma radiation)	GTK	Nation wide	No updates	Raster	Open data	low		classification of peat layer	Moisture content of soil	GTK[19]

[17] www.maanmittauslaitos.fi/en/digituotteet/laser-scanning-data

[18] www.maanmittauslaitos.fi/

[19] hakku.gtk.fi/en/locations/search

Elevation and surface models

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Elevation model 2m	NLS	Partial	-	Raster, 2 m	Open data	Accuracy (z 0.3 m)	High	Terrain visualization	Operations planning	NLS[20]
Elevation model 10m	NLS	Nationwide	-	Raster, 10 m	Open data		High	Terrain visualization		NLS[21]
Canopy Height Model (CHM)	ALS Producers	Partial	-	Raster	Restricted		High	Forest biomass visualization	Estimates of terrain trafficability, operations planning	
Digital Elevation Model (DEM)/Digital Surface Model (DSM)	ALS producers	Partial	-	Raster	Restricted		High	Terrain visualization	Estimates of bearing capacity, operations planning, operator tutoring	
Shaded	NLS	Nationwi	-	Raster	Open data		High	Terrain	Operations	www.maanmittauslait

[20] www.maanmittauslaitos.fi/en/digituotteet/elevation-model-2-m

[21] www.maanmittauslaitos.fi/en/digituotteet/elevation-model-10-m

relief		de						visualizatio n	planning, operator tutoring	os.fi

Map data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Basic map raster	NLS	Nationwide	Varying between types of information	Raster, Map scale 1:10000	Open data	Good	High	Visual planning of operations, base maps		NLS[22]
Topographic database	NLS	Nationwide	Varying between types of information	Vector, Map scale 1:10000	Open data	Good	High		Gis analyses, Operations planning	NLS[23]
Soil map/ Superficial deposits	GTK	Scales and coverage: 1:20000; partial 1:50000; partial 1:200000; nationwide	N.a.	Vector	Open data	Varying accuracy	High	Soil classification	Estimate of trafficability	GTK[24]
RLGis data	SMK, MH	Partial	N.a.	Vector	Cost/ Restricted	-	-	Estimates of erosion potential	Estimate of trafficability	-
Static trafficability maps	SMK/Arbonaut	Partial	N.a.	Raster	Open data	Varying accuracy	-	Estimate of trafficability, soil classification		

[22] http://www.maanmittauslaitos.fi/en/digituotteet/basic-map-raster
[23] www.maanmittauslaitos.fi/en/digituotteet/topographic-database
[24] hakku.gtk.fi/en/locations/search

Thematic databases

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Land use plans; general master/ comprehensive plan	Provinces, municipalities, state, Syke	Planned areas, Partial	On demand	Vector	Open data	Varying	High	Operations planning		SYKE[25]
Nature conservation areas	Syke, MH	Nationwide	Biannual	Vector	Open data		High	Operations planning, operator tutoring		SYKE[26]
Groundwater areas	Syke	Nationwide	Biannual	Vector	Open data		High	Operations planning		SYKE[27]
Cadastral map	NLS	Nationwide	Up-to-date	Vector/Raster	Open data		High	Operations planning,		NLS[28]

[25] http://www.syke.fi/en-US/Open_information/Spatial_datasets
[26] http://www.syke.fi/en-US/Open_information/Spatial_datasets

[27] http://www.syke.fi/en-US/Open_information/Spatial_datasets

[28] http://www.maanmittauslaitos.fi/en/digituotteet/cadastral-index-map

								operator tutoring		
Digiroad -	Finnish Transport Agency	Nationwide	Monthly	Vector; road geometry, road condition parameters	Open data	varying	High	Transport planning, forest road maintenance and construction	Transport optimization	http://www.liikennevirasto.fi/

In-situ data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Field inventory data of forest planning	SMK, MH, forest owning companies, forest management associations, OTSO	Partial	On demand	Vector	Restricted	Varying	Medium to high	Forest management planning		
Forest machine data	Forest machine entrepreneurs, machine manufacturers	Operation areas	Continuous, daily basis	HPR, PRI, STM etc.	Restricted	Varying	StanForD, high; post-harvest quality and other data, varying	Data for forest operations, logistics, control of bucking, payment etc., industrial processes, harvesting quality monitoring	Forest management planning, silviculture operation planning, , estimates of trafficability	
Photography based	Trestima, SME	Inventory areas	On demand	Vector	Restricted, cost	Varying accurac	-	Forest inventory	harvesting quality	Trestima [29]

[29] https://www.trestima.com/products_en/#trestima

forest inventory						y		and planning	monitoring	
Timber receiving station data -	Forest industry	All industrial timber	Daily basis	Measurement and quality metrics	Restricted	High	High	Control of timber quality and measurement, production management, measurement for timber trade	Bucking and supply chain optimisation and quality control	

Meteorological data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Weather data	FMI	Nationwide/ ~400 points	Daily	Vector	Open data	High	High	Operations planning, trafficability estimation	Weather sensitive trafficability estimation	en.ilmatieteenlaitos.fi/open-data
Interpolated weather grid	FMI	Nationwide	Daily	Vector, 10km grid	Restricted	Medium	High	Operations planning, trafficability estimation	Weather sensitive trafficability estimation	
Weather forecasts	FMI	Nationwide	Daily	-	Open/cost data	Varying	High	Operations planning, trafficability estimation	Weather sensitive trafficability estimation	en.ilmatieteenlaitos.fi/open-data
Hydrological observation data	Syke	Nationwide	Daily	-	Open data	High	High	-	Estimation of moisture content	www.syke.fi/en-US/Open_information/Open_web_services
Forest fire index	FMI	Nationwide	Daily during summer	Vector, 10km grid	Restricted	Variable	High	Forest management planning		http://en.ilmatieteenlaitos.fi/forest-fire-index
Soil frost observation	Syke	Nationwide	Daily	-	Open data	High	High		Forest operations planning	www.syke.fi/en-US/Open_information/Open_web_services

Forest estimates

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
ALS forest estimates	SMK, MH, Forest owning companies, SME	Partial	On demand	Vector, 16*16m grid	Restricted	Varying	High	Strategic planning	Raster data of NIPF owners open soon, planning of wood buying and procurement	www.metsaan.fi/
Forest resource data for stands	SMK, MH, forest owning companies, OTSO, MHY, SME	75 % cover NIPF, 100 % other	Max. 10 years	Vector, compartment based	Restricted	Varying	High	Operations planning, strategic planning		www.metsaan.fi/
Nationwide Forest Inventory (NFI)	Luke	Whole country	Continuous	Forest statistics of all estimates, thematic maps	Varying (open statistics data and restricted raw data)	Good	High	Strategic and scenario planning		www.luke.fi/en

Statistics

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Forest sector statistics	Luke	Nationwide	Monthly	Publications, database of estimates	Free	High	High	Strategic and scenario planning		www.luke.fi/en
Industry statistics	Forest companies	Company specific coverage	Daily-basis	m, m3, tonne, € etc.	Company specific, restricted	High	High	Operational and strategic planning		

France

Introduction

In the case of France, expectations towards EFFORTE are very much focused on the better understanding of forest soil trafficability and the different ways to prevent potential damages which hinder both productivity and sustainability. Strategic and operational prevention, supported by knowledge-based planning, have a significant role to play in the way forest practitioner's deal with soil trafficability. But decision-support system capable of integrating the multifactorial dimension of soil trafficability are for now still very limited and often experience-based. However, integration and joint exploitation of different BIG DATA sources could reasonably contribute to filling that gap.

The here-under described data sources are currently considered as relevant pieces for the prototype tool under development by French partners in EFFORTE.

Remote sensing

Remote sensing data sources can provide important context for the planning of forest operations.

Data source	Producer	Coverage	Update frequency	Delivery format	Access-ibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
CORINE Land cover	Ministry for sustain-able dev. & energy	Nationwide (+ 38 other European countries)	Latest 2012	Vector	Free	Good Smallest mapped unit = 25ha	High European nomenclature (44 levels including 3 forest levels)			http://www.statistiques.developpement-durable.gouv.fr/clc/fichiers/
Ortho-photo BD FORET	IGN	Nationwide	Version 1 until 2006,	Vector	Free for resear	Version 1: min.	High New			http://professionnels.ign.fr/ign/configure/74096/purchase

(nomenclature based on 32 itrems)			since then version 2 with first Nation-wide coverage in 2016		ch, other wise costly	mapping unit 2.25 ha, version 2: min. mapping unit 0.5 ha	method since 2007			
Orthophoto BD ORTHO	IGN	Nationwide	Every 3 years	Raster	Free for research, other wise costly	50 cm of resolution	High			http://professionnels.ign.fr/bdortho-50cm
LIDAR data	IGN or sub-contractor	Limited to small experiments 1pt/m²	Project-based	Point cloud	For research only	Good	High		Drainage, damage assessment	
Very High Resolu-tion data SPOT 6/7	Spot Image	Nationwide	Data acquisition period : from 3/02/2014 to 1/11/2014 Other period	1.5 – 6m	Free for research, other wise costly		High		Damage assessment	http://www.intelligence-airbusds.com/fr/870-spot-6-7

			on demand and costly							
Very High Resolu-tion data RapidEye	GeoSys	Nationwide	Data acquisition period : from 05/2010 to 10/2010 Other period on demand and costly	5m	Free for research, other wise costly		High			https://datasourcing.geosys.com/rapideye-partenaire-exclusif-pour-la-france/

Elevation and surface models

Elevation and surface data sources can provide information on topography with different levels of accuracy. The Nationwide elevation reference cover in France is described below. Lidar Nationwide cover is not yet available in France even if they would be preferred for questions of accuracy.

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Digital Terrestrial Model **BD ALTI**	IGN	Nation-wide	Only if important relief evolution	Raster Special resolution down to 25m	Free, down to 75m	Good	High	Drainage	Drainage	http://professionnels.ign.fr/sites/default/files/DC-BDALTI_2-0.pdf

Map data

The four here-down described data sources about soil, geology and hydrology are provided as Nationwide maps in France. They all contain some of the factors which define soil bearing capacity as a function of soil type and water content. Moreover, public forest boundaries are also delivered as maps by the State forest office. As such they provide significant information on the ownership type and the special regulation that may apply regarding traffic by forest machines.

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Soils **BDGSF**	GIS SOL	Nation-wide	1985 (No update)	Vector 1:1000000	Cost	For global analysis only	High European standard BDGSE	None but research on soil compaction and trafficability at nationwide level		http://acklins.orleans.inra.fr/programme/bdgsf/contenu.php http://www.afes.fr/afes/egs/EGS_2_3_JAMAGNE.pdf
Geology BD CHARM	BRGM	Nation-wide	No update since 2013	1:50000	Cost	Good	High	Indirect		http://www.brgm.fr/sites/default/files/plaquette_charm_50_2015.pdf
Public forest boundaries	ONF	Nation-wide	Only in case of change	Vector	Free	Good	High			http://carmen.carmencarto.fr/105/ONF_Forets.map

Contour forêt publiques										
Water catchment & hydraulogy **BD CARTHAGE**	Agence de l'Eau	Nation-wide	Annual	Vector, deca-metric accuracy	Free	Good	High	Research on soil compaction and trafficability		http://services.sandre.eaufrance.fr/telechargement/geo/ETH/BDCarthage/FXX/2014/

Thematic databases

Administrative boundaries available as thematic databases can be useful to apply special regulations (in regard to traffic by forest machines or forest operations in general) to the local area they relate to.

Data source	Producer	Coverage	Update frequency	Delivery format	Access-ibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
GEOFLA Administrative boundaries	IGN	Nationwide, municipalities	Annual	Vector 1:1000000	Free	High	High	Indirect		http://professionnels.ign.fr/sites/default/files/DC_GEOFLA_2-2.pdf
BD CARTO Administrative boundaries, road network, land-use map…	IGN	Nationwide	Every 1 to 5 years depend-ing on items	Vector 1:50000	Free for research, other-wise costly	High	High	Indirect		http://professionnels.ign.fr/sites/default/files/DC_BDCARTO_3-2.pdf
BD TOPO Administrative boundaries, road network, land-use map…	IGN	Nationwide	Every year (road and rail layers) to 5 years (other layers)	Vector 3D, metric accuracy	Free for research, other-wise costly	High	High	Indirect		http://professionnels.ign.fr/sites/default/files/DC_BDTOPO_2-2.pdf
Environ-	INPN	Nationwi	Annual	Vector	Free	Vari-	High			https://inpn.mnhn.fr/telechargement/carte

mental areas (Natura 2000…)		de		(from 1:10000 to 1:100000)		able				s-et-information-geographique

In-situ data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
In-situ soil fertility description **BDAT**	INRA	France Whenever a soil sample is sent to INRA to be analyzed	~ 250000 entries/year	Online access geographic interface	Free	High	High, in accordance to testing protocol			https://www.gissol.fr/donnees/tableaux-de-donnees/donnees-de-la-bdat-3028

Meteorological data

Meteorological data source include information about past, present and future water inputs to forest soils as well as the weather condition which play a role in natural evaporation. Considering the importance of water content on soil bearing capacity, this information represents a critical source for strategic and operational planning.

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Weather stations (locations)	Météo France	Nation-wide		Point	Free	Good	High			https://publitheque.meteo.fr/okapi/accueil/okapiWebPubli/index.jsp
Weather records - **Rainfall** - **evapotranspiration**	Météo France	Nation-wide	Hourly	Points	Costs depend on frequency	Good	High			https://publitheque.meteo.fr/okapi/accueil/okapiWebPubli/index.jsp
Weather forecast **ARPEGE**	Météo France and Euro-pean Center CEPMMT	Nation-wide	Every 6 hours (forecast up to 102 hours)	Raster 7.5km reso-lution						http://www.umr-cnrm.fr/spip.php?article121

Forest estimates

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Existing applications in planning and wood supply	Potential applications in planning and wood supply	Reference
Nation-wide forest inventory	IGN	Nation-wide	Rotating data collection over 1/10 of the nation-wide plots	Points	Mostly free	Good	High	Standing resources availability, research on soil sensibility		http://inventaire-forestier.ign.fr/spip/spip.php?article532

Statistics

There are no statistics available in France concerning soil trafficability.

Additional information on the use of these data sources in soil trafficability research

BDGSF (Nationwide soil map), with its texture classification, can be used in small-scale map to represent soil sensitivity towards compaction.
Soil classification (type of soil, texture, pseudogley oncoming depth, stone rate) at each plots of the Nationwide Forest Inventory can also be extrapolated to generate a small-scale map of soil compaction sensibility.
BD CHARM (Nationwide geological layer) with BD ALTI (digital terrestrial model) is used in France by BRGM to calculate IDPR index. This index represents soil capability for runoff and infiltration. However, IDPR index from BRGM is not easily available by other companies. That's why it's also possible to asses IDPR index with a model based on BD ALTI (digital terrestrial model) and BD CARTHAGE (Hygrology model).
By combining compaction soil sensibility map and weather records, it could be possible to assess soil dynamic sensibility. But scales of data only allow a strategic use and not an operational use. Indeed, it's not possible, with today's data, to predict compaction soil sensibility in a forest plot or a logging area.

4. Potential future Big Data sources

4.1. Overview

The potential future Big Data sources are based on existing knowledge of data sources and most of them exist in research studies or with increased availability in the coming years. They hold the potential to significantly support the models or decision support to be developed to support the EFFORTE objectives and much more related to forest planning and wood supply. It is foreseen that the forest industry will require even more details on the raw material which might be possible to achieve with listed potential future Big Data sources.

4.2. Sweden

Introduction

The most critical issue in Sweden for the future is the fees on data from Lantmateriet. The Agency wants to be able to offer all data free of charge, but require additional funding from the government. Most likely this will happen in the coming years and then we foresee even more models and decision support developed and marketed by SMEs. Another critical issue is related to the forest estimates from lidar data where both agencies and companies are trying to find funding for a new coverage and updated forest estimates.

In the lists some current available data sources are listed, but as they are not yet used in operations they are candidates for potential future Big Data sources.

Remote sensing

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Lidar from UAV	Service providers	Partly	On request	More than 300 000 points/s	Cost		Low		e.g leica-geosystems.com
Mobile laser scanners	Service providers	Partly	On request	More than 300 000 points/s	Cost		Low	Detailed information on stems prior or	e.g www.cowi.se www.terratec.se

								during forest operations	
Airborne lidar data (possibly new campaign)	Lantmateriet	Nationwide	1 coverage	0.5-5 points/m^2	Free		High	Forest estimates, mapping of environmental and cultural considerations	www.lantmateriet.se
VHR satellite data	Planet Labs	Nationwide	On request	1 m	Cost		Medium	Detailed planning in production and considerations, damage assessment, monitoring, forest estimates	https://www.planet.com/
Ground-based radar	Service providers	Partly	On request	Varying	Cost		Low	Details on soils in trafficability maps	Service providers or www.sgu.se
Airborne multispectral lidar data	Service providers	Partly/ nationwide	On request	2-15 points/m^2	Cost		Medium	Forest estimates including tree species mapping	e.g. www.terratec.se
Mobile phone images or videos	Service providers, forest companies	Partly	On request	Vary	Free, restricted		Low	Forest estimates including tree species mapping	e.g. Trestima Katam

Elevation and surface models

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
DSM from radar satellite	Service providers	Nationwide	Every 2-3 weeks	Raster	Cost		Low	Tree height, timber volume, coverage density estimations in forest	e.g. EADS[30]
DSM from optical satellite	Service providers	Nationwide	Every 2-3 weeks	Raster	Cost		Low	Tree height, timber volume, coverage density estimations in forest	e.g. www.metria.se

[30] http://www.intelligence-airbusds.com/terrasar-x/

Map data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
CadasterEnv (detailed land cover mapping)	Swedish Environmental Protection Agency	Nationwide	3-5 years	Raster, 10 m	Free	Thematic accuracy 80%	Medium	Tree species mapping and environmental considerations	www.cadasterenv.se/

In-situ data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Crowdsourcing	General public	Partly	n.a.	WMS (most likely)	Free	Low	Low	Road conditions, environmental or cultural considerations	e.g. https://crowdmap.com/welcome

Meteorological data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Hydrological models, HYPE (weather data, frozen ground, wind etc.)	SMHI	World (regional on water catchment areas)	1-10 days forecast	Database	Cost	Catchment area	High	Trafficability maps, decision support	www.smhi.se

Forest estimates

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Skogliga grunddata (lidar derived forest estimates) Potential update	Swedish Forest Agency	Nationwide	10 years	12,5 m raster	Free	< 20 % Mean average error in estimates	Medium	Operational planning, timber purchase support environmental considerations	www.skogsstyrelsen.se
Local forest estimates from mobile phone apps	Service providers, forest companies	Partly	Continuous	Vary	Cost (restricted)	n.a.	Low	Timber purchase support	

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Forest operations statistics in real time	Forest machine companies	Global	Continuous, (Internet of things)		Cost	High	High	Wood supply support	Forest machine companies
Forest transportation statistics in real time	Truck companies	Global	Continuous, (Internet of things)		Cost	High	High	Wood supply support	Truck companies
Information from Forest machine and truck conditions in real time	Forest machine companies	Global	Continuous, (Internet of things)		Cost	High	High	Sustainable forest operations (without interruptions)	Forest machine companies

4.3. Finland

Introduction

Various novel Big Data applications are about to come soon. One important aspect for developing new solutions is to make data fusion of available Big Data. Other important drivers for this development are i) automation and measuring technology (robotics, machine vision), ii) information networks and data terminal equipment, iii) analysing, transferring and storing capacity of information, iv) analysing methods of Big Data and information standards and v) models produced in the forestry industry, as example.

Modern measurement systems, like mobile laser scanning systems will be installed to forest machines to collect enormous amounts of high quality data for various end use possibilities. Another rapidly developing data collection method in the future is crowdsourcing. For example, with this method forest professionals and owners can produce a lot of different kind of data with mobile applications.

Data fusion, combining multiple data sources can create valuable new data, more accurate estimations and intelligent applications. For example, dynamic trafficability map can combine soil data, weather data, forest estimates and digital elevation models.

Remote sensing

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
UAV photos/laser scanning	SME (small and medium sized entrepreneurs), NIPF owners, Forest companies, MH, SMK	case specific, small scale	case specific	raw data and estimates	cost	High	High	forest estimates, tree mapping, quality estimates and monitoring	
Mobile laser scanning	Logging operators	Small scale	daily-basis	LAZ, Vector	Cost	High	High	Forest estimates, tree map	

(MLS)								data for accurate forest estimations, quality monitoring, operator tutoring		
Terrestrial laser scanning (TLS)	operators related to forestry	Small scale	on demand	LAZ, Vector	Cost	High	High	Forest estimates, quality monitoring		
VHR satellite data	Planet Labs	Nationwide	On request	1 m	Cost		Medium	Detailed planning in production and considerations, damage assessment, monitoring, Forest estimates	https://www.planet.com/	

Map data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Electric lines, Ground cables	-	Nationwide	-	-	Open	High	High	Security planning for operations	

Thematic databases

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Soil moisture estimates/database	-	Nationwide	Daily-basis	-	Cost	-	-	Operation planning (e.g. wetness map)	
Weather sensitive soil bearing capacity estimates/database	-	Nationwide	Daily-basis	-	Cost	-	-	Operation planning, operator tutoring	
Dynamic road condition database	Transport entrepreneurs, SME	Nationwide	Daily-basis	-	Open	Varying	-	Planning of daily wood transports, operator tutoring, navigating apps	

In-situ data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Soil database	Forest machine entrepreneurs, SME	Most of commercial forests	Daily-basis	Soil and terrain feature estimates database	Open	Varying	Varying	Forest operations planning, terrain trafficability, accessibility, operator tutoring	
Stem dimensions and quality database	Forest machine entrepreneurs, forest companies, SME	Nationwide	Daily-basis	Tree quality estimates in spatial database	Restricted	Varying	High	Planning of wood pricing and purchase, planning of bucking pre-orders	
CAN-bus data of harvesters for assessing trafficability	Forest machine entrepreneurs	In most commercial forests	Daily-basis	Point/ vector data and attribute	Restricted	Varying	-	Routing of forwarder's traffic	
Post-harvest quality database	Forest machine entrepreneurs, forest companies	In most of commercial forests	Daily-basis	Post-harvest quality estimates database – several possible parameters (point/ vector and	Restricted	Varying	High	Post-harvest quality monitoring, operator tutoring, operations planning	

				attribute)					
Harvester database	Forest machine entrepreneurs, forest companies, SME	In most of commercial forests	Daily-basis	-	Restricted	Varying	High	Planning of wood pricing and purchase, planning of bucking preorders, updating forest resource data, reference data for remote sensing	

4.4. France

Introduction

Considering the French focus on soil trafficability only a limited number of potential future Big Data sources are here-down considered.

The most promising and realistic future data sources in France are based on lidar technology. The exponential growth of lidar sensors due to the development of autonomous vehicle market is going to benefit forest applications: miniaturization, lower cost of sensors and economy of scale, easier implementation and more.

Remote sensing

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Lidar from UAV	Service providers	Partly	On request		Cost	High	High	Rutting evolution in an area	
Mobile laser scanners	Service providers	Partly	On request		Cost	High	High		
Airborne lidar data	Service providers or IGN ?	Nationwide	Regular campaign	1 point/m^2	Cost	High	High		
Terrestrial laser scanning	Forest operator or IT company	Local	On request	Vector, LAZ	Cost	High	High	Forest estimates	

Elevation and surface models

If the above-mentioned remote-sensing-based data source were to become available, more accurate surface and elevation models would be accessible.

Map data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Water Catchment & hydraulogy BD TOPAGE	IGN - ONEMA	Nationwide After first Experimental stage	Not specified	Vector Metric accuracy	Free for research, otherwise costly	High		Drainage	
Experience-based maps about year-round soil trafficability	Forest managers such as large public/private organisations	Regional	As often as the consensus would need to be debated again	Raster	Free	Medium	Medium if relying on Nationwide initiative	Prevention when providing logging instructions	A few initiatives do exist in France especially at regional level when state forest office is confronted to trafficability issues which are hard to assess

In-situ data

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Harvester data including routing	Each machine or company	Stand level (machine)	Daily per machine or once the operation is finished	.hpr (.xml) .pri (ascii)	Restricted	Medium	High if StanForD 2010 is implemented	Driving patterns within the stand	
Forwarder data including routing	Each machine or company	Stand level (machine)	Daily per machine or once the operation is finished	.hpr (.xml) .pri (ascii)	Restricted	Medium	High if StanForD 2010 is implemented	Driving patterns and transported volumes	

Meteorological data

One possible improvement of the meteorological data sources could result from the increase and distribution of weather stations, for data to be collected in more locations and more frequently in forest areas.

Statistics

Data source	Producer	Coverage	Update frequency	Delivery format	Accessibility	Data quality	Level of standardization	Potential applications in planning and wood supply	Reference
Logging observation	Public association	Nationwide	Annual or in real time	Excel	Free	Forest plot	High		

5. SWOT analysis

5.1. Overview

The traditional SWOT analysis contains strengths and weaknesses with internal origin and opportunities and threats with external origin. Our analyses define the internal origin as features of the dataset and external origin as the opportunities and threats for use in the forestry sector and especially for use in models and refined datasets ready to use in forest planning and operations.

Each country section start with a general SWOT analysis concerning Big Databases in general for forest planning and wood supply, but with a focus on trafficability and decision support to forest operations.

The SWOT analysis is based on the most (maximum 5) important Big databases country-wise from each category. This selection is performed to highlight the present and future databases of greatest importance for the forestry sector.

5.2. Sweden

Introduction

A general SWOT analysis of Big Databases in forest planning and wood supply show the challenges and opportunities in Sweden:

Strengths - Many data sources available and easy accessible - Some decision support on trafficability implemented – acceptance by users - Focus in forestry in decision support – interest to further develop and implement	**Weakness** - Relatively high costs on detailed Big Data sources - GNSS in forest machines can be rather poor - Soil maps and Met office data not developed for forest operations
Opportunities - Many Big Data sources coming - Improved tools for Big Data Analytics - Crowdsourcing may support data collection - Support from investors and research community on Big Data data and processing methods	**Threats** - Too much variation in data quality reduce the reliability of the decision support or models - Models that require data with restricted access (such as private data) may not be developed although data quality is high - New technology may not deliver useful results

Existing

Datasource	Strengths	Weaknesses	Opportunities	Threats
Orthophoto	Part of Nationwide programme, established technique, large coverage	As part of Nationwide programme only every second acquisition is during vegetation period, not linked to administrative borders	Image analysis for improved decision support	Global or EC programmes with free image data
Historic Orthophoto (Year 1930-1990)	Nationwide coverage, time series	Varying technical quality	Time stamps relating to forest conditions and change	Lack of interest in forest applications
High resolution optical satellite data	EC programme secure continuity, short revisiting periods	Large data volumes, Big Data challenges, relatively low resolution for detailed forest services	On-line processing, large investments in services	Nationwide campaigns that reduce costs for alternative data
Airborne lidar data	Detailed data and new technology to be introduced commercially (photon scanners)	Relatively expensive, Big Data challenges	Repeated mapping of forest estimates, forest growth mapping	Nationwide campaigns that reduce costs for alternative data
Surface model from orthophotos	Frequent update interval, nationwide coverage, height estimates almost as reliable as lidar data	Point cloud insufficient for detailed reliable forest estimates	Image analysis for improved decision support, forest growth mapping,	Nationwide campaigns that reduce costs for alternative data
Grid 2+	Nationwide coverage, detailed dataset	Not yet complete	Image analysis for improved decision support	Private (or EC) campaigns that reduce costs for alternative data
Depth to water maps	Nationwide coverage, widely used, easy to interpret	Several different versions, varying accuracy depending on soil type, climate and artificial structures	Image analysis for improved decision support	Changes in climate variables
Soil maps	Nationwide, detailed in regions	Content vary, “too detailed”, not fully standardized	Improved decision support in models, soil information significant in forest operations and information is stable	Contribution in forest applications unknown
FMIS	Nationwide and detailed information content	Incomplete with positioning errors	To be part of improved decision support in forest operations	Crowdsourcing more relevant
NVDB	Standardized and nationwide	Incomplete on private roads	To be part of improved decision	Competitors, like Open street map,

			support in forest operations	more easily updated
Green infrastructure	Nationwide, based on best available sources	No updates including known changes, unknown in forestry applications	To be part of improved decision support in forest operations	More up to date and accurate datasets
Harvester data	Standardized and accurate with high resolution	Not implemented in all harvesters yet, company databases under construction	To be part of improved decision support in forest operations	Incomplete registration
Forest stand database	Well established, many parameters	Varying quality, often outdated	Further refined to support decision support, Reference data	Grid-based forest estimates
Climate predictions	Nationwide with calculated with scientific methods	Rather course resolution	To be part of improved decision support in forest operations	Information content to be investigated
Lidar derived forest estimates	Reliable forest estimations, potentially nationwide	Rather costly, requires in-situ data as reference	Grid-based forest estimates to replace forest stand databases	Nationwide campaigns that reduce costs for alternative data
Tree height maps	Reliable detailed forest height estimations, potentially nationwide	Underestimation of tree height, quickly outdated	Further refined to support decision support, Reference data	Better accuracy in competing products

Potential future data sources

Datasource	Strengths	Weaknesses	Opportunities	Threats
VHR satellite data[31]	High spectral and temporal quality, possible to order over area of interest	Relatively high cost	Forest estimates with similar quality as airborne lidar estimates	Nationwide campaigns that reduce costs for alternative data (airborne lidar)
Lidar from UAV	Detailed and local data collection, rapid data collection	Unknown legal status, challenging data processing	Very detailed and reliable forest information	Price and possible low range
Mobile laser scanners	Detailed and local data collection	Challenging data processing, unknown information content	Very detailed and reliable forest information	Other cheaper techniques, like mobile phones apps
Airborne multi-spectral lidar data	Map tree species and detailed forest structure	Challenges with large coverage	Very detailed and reliable forest information	Other cheaper techniques, like mobile phones apps
DSM from optical satellite	Detailed and frequent	Relatively high cost, not fully investigated	Very detailed and reliable forest information	Nationwide campaigns that reduce costs for alternative data (orthophotos)
CadasterEnv	Nationwide based on best available data sources	Thematic product, generalized	Possible detailed and reliable forest stand information	Information content to be investigated
Crowdsourcing	Cheap and local	Quality assurance, incomplete (most likely)	Frequently updated information for further analyses	Nationwide datasets from authorities
HYPE	Scientific method with nationwide coverage	Low resolution (sub-catchment areas)	To include seasonal variation	Information content to be investigated
Updated lidar forest estimates	Reliable forest estimations, potentially nationwide	Rather costly, requires in-situ data as reference	Grid-based forest estimates to replace forest stand databases	Nationwide campaigns that reduce costs for alternative data
Local forest estimates from Mobile phone apps	Cheap, rapid and local	Quality assurance, incomplete (most likely)	Frequently updated information for further analyses	Grid-based forest estimates
Forest operations statistics in real time	Standardized, detailed information on wood supply, progress of	Coverage of data transmission services	Up-to-date wood supply statistics to industry	Data confidentiality

[31] Planet labs concept

	harvesting			
Forest transportation statistics in real time	Standardized, detailed information on wood supply, progress of transportation	Coverage of data transmission services	Up-to-date wood supply statistics to industry	Data confidentiality, legislation
Information from Forest machine and truck conditions in real time	Standardized, detailed information on machine status	Coverage of data transmission services	Improved machine service status and monitoring	Data confidentiality, "society of monitoring"

5.3. Finland

Introduction

The general SWOT-analysis of Finnish forestry Big Data sources is expressed at first. More detailed analyses of the most important data sources are presented in the following tables.

Strengths - A lot of open data is available already in Finland - Forest estimates with high coverage and quality are mainly already available - Mapping has been used a lot for various forestry activities in nationwide, thus offering a full coverage	**Weaknesses** - Present soil maps are inaccurate to precision forestry. In addition, the coverage of the most accurate soil map is only partial - Positioning accuracy (GNSS) can be low
Opportunities - A lot of weather data is available which enables modeling of soil moisture conditions and trafficability - Data fusion creates various opportunities in utilization of Big Data, - Forest machines produce already a lot of data and data collection by machines can be still extended with sensor technology - Crowdsourcing offers continuous and wide-scale data collection - Analyzing of digital photos can open possibilities to create data for different purposes in forestry	**Threats** - Ownership of data can have an influence on larger data collection and data providing - New data collection can be expensive and cause problems in funding, - Allocating of data collection expenses can be complicated - Several technologies competing each other can form risks to application developers

Existing

Datasource	Strengths	Weaknesses	Opportunities	Threats
ALS	Versatility of data use, high coverage-data collection ratio, Fast method	Detecting tree species, lower canopy trees and field layer, quality variation, weather dependent	Usability to combine and update data with other data sources, possibilities to increase data use for other purposes, higher resolution and improved data content (tree quality, dimensions)	
Forest machine data	Mainly standardized data, large volume of data, data for updating forest resource data, reference for remote sensing data	Restrictions of data use, scattered data collecting and handling	Possibility to increase the use of data a lot; operator tutoring, post-harvesting quality, forest management planning, bucking quality	Ownership of data, privacy protection, availability of data
Forest resource data	High coverage, versatility of data and data use, expected to be open data soon	Variation of quality, restrictions in use of some part of data, expensive field measurements	Usability and accuracy will be increased when combined to forest machine data, ALS, other data	Ownership of data, privacy protection, availability of data
Nationwide forest Inventory data	High coverage, multi-source data, opportunities for scenario analyses and trends, versatile open statistics	Not for operational use, rather expensive collection procedures, scale of metrics	Possibility to combine with other data sources	Funding of inventory
Soil map/superficial deposit map	High coverage, importance of data in forestry	Low accuracy, high spatial variation of soil type, no development anymore, no updates	Possibility to combine with other data sources	Low accuracy prohibits the utilization
Basic map raster	Coverage, well known	Part of features are inaccurate	Accuracy improvement in topography	
Weather data and forecasts	Accurate history database, short-term accuracy	Spatial variation of weather phenomenon, long-term accuracy, scale of metrics (10x10km)	Possibility to combine with other data sources, dynamic trafficability mapping	Inaccurate spatial weather database

Potential future data sources

Datasource	Strengths	Weaknesses	Opportunities	Threats
MLS	High accuracy, data collection combined to other operations	Expensive, sensitive for field conditions, data management and data transfer is challenging	Possibilities to update and improve other data, improved stem quality scanning, bucking, post-harvest stand data	Ownership of data
Dynamic road conditions database	Improved data of road conditions	Variation on data, limited coverage	Better road quality and maintenance, higher operational efficiency	Availability of input data
Weather sensitive trafficability database	Improved timing of forest operations, site selection, operational efficiency, decreased environmental impacts	Sensitive for accuracy of input data	Decreased seasonal variation and its effects, improved acceptability and operational efficiency	Availability of input data
Stem quality database	Accurate stand and stem quality data	Variation in accuracy, compiling final data, quality availability	Improved planning of wood pricing and purchase, planning of bucking pre-orders, silvicultural treatments	Ownership of data
Precise satellite imaging	High spectral and temporal quality, possible to order over area of interest, short scanning interval	Relatively high cost in some sources	Monitoring of forests (storms, cuttings etc.), combined platform for satellite databases	
Post-harvest quality database	Data coverage, instant quality reports, standardized data	Accuracy of some estimates	Improved acceptability, improved forest management quality, operator feedback	
Soil database	Improved quality and accuracy of soil data, reference data for other estimates, data collection combined to other operations	Variation in data quality, low update frequency	Improved accuracy and quality in planning of forest operations, better trafficability maps, decreased seasonal variation	Funding, demanding development of soil identification methods

5.4. France

Introduction

SWOT analysis relates to Big Data sources which are available today for trafficability use. Proposals on how critical lack of data could be overcome are also described.

Strengths	**Weaknesses**
- Need for new solutions. Forest work is heavily disturbed by high seasonal variation. Forest managers and logging	- Scales of available data are too small - Lack of information on logging activity in a large-scale geographic repository. It's

operators need a better knowledge and prediction to secure their license to practice. - Over the past few years, low machine utilization rate and costs due to skid trails rehabilitation facilitated changes of awareness/mentalities which are necessary to look for solutions through BIG DATA and decision-support system for better planning and monitoring.	necessary to measure and monitor harvesting and forwarding impacts (for instance, monitoring of rut impact in permanent skid trails with regular lidar data)
Opportunities - To create an observatory of logging operations (centralize information on harvesting /forwarding operations; database of logging administrative declarations; confidential and anonymous process) - To create regular lidar campaign in a forest area or at a Nationwide level or to use very high resolution satellite data - To measure changes over time on permanent skid trails (damage assessment) - To build a knowledge data base on interactions between logging operations / weather records / ruts with key indicators. - To develop a model for trafficability and damage prediction	**Threats** - Usual stakeholders' reluctance to share their data in a centralized data base - Difficulties related to change management - Computing power to be developed

For existing and potential future data sources, a few critical examples are listed below.

Existing

Datasource	Strengths	Weaknesses	Opportunities	Threats
VHR satellite imagery	Automatic data-collection, on demand	Too costly for operational use	Very short period / frequent up-date	Data use in private forest (70 % in France) Regulation concerning private digital data (data protection, owners permission)
Historic Orthophoto	Time series, nationwide reference frame	Varying technical quality	To be combined with terrestrial digital model from lidar	Process not user friendly

Potential future data sources

Datasource	Strengths	Weaknesses	Opportunities	Threats
UAV Lidar data	Easy to collect, very flexible, well adapted to a small area or medium area survey	Automatic acquisition more difficult than satellite imagery, technology still needs further development, miniaturization of lidar technology to be compatible with UAV	Flexibility (easy to carry out in a new area)	Constant evolution in UAV, regulation (more constraints to be expected in the future), processing methods to be improved

6. Conclusions and outlook

Based on the country reports some common key datasets are identified for use in R&D to develop operational methods and tools in forestry for the EFFORTE objectives. They are not yet available in detail and nationwide in all three countries, but may be there soon.

Some dataset requires further development in terms of Big Data analytics and Internet of Things (IoT) to automate data collection. The inclusion of machine data may be fully automated and this is needed to crap the full potential of these datasets. A manual data handling will require too much resources in the long run.

The level of maturity and European dimension describes both the availability in European countries and the status in operational implementation. Further efforts are generally needed to exploit how the datasets may be utilized in forest operations and for the objectives in the EFFORTE project. Applications relating on forest machine data, VHR imagery and mobile laser scanning are candidates to potential future datasets that requires with both additional R&D in data collection and how to be utilized as decision support in forest operations.

Datasource	Justification	Level of maturity and European dimension
Detailed digital elevation model[32]	A detailed digital elevation model is the bases for most terrain trafficability maps and a key parameter to avoid soil disturbances, accelerate machine mobility and assess persistence of soil compaction and rutting.	Available today in many European countries.
Lidar based forest estimates	Forest estimates, such as timber volume or diameter, are important in forest planning and wood supply and also key factors in harvesting operations and detailed production planning, e.g. in terms of rutting. Lidar based forest estimates may be based on the same lidar collection as the detailed digital elevation model. New techniques in data collection, e.g. photon scanning will most likely make it faster and cheaper.	Available today in some European countries.
Soil maps	Soil maps may be further utilized in forest operations and trafficability mapping	Available today in many European countries, but with varying spatial resolution that need further studies for implementation in forest planning.
Weather data and models	Weather maps and models plays an important role for the bearing capacity and could play an important role for more detailed planning of forest operations to avoid soil disturbances and for improved productivity in forest operations.	Available today in all European countries, but not included in forest operations models or decision support.
Road databases	For logistics, the road databases play an important role to plan wood allocation to industry, but also in getting access to planned forest to cut where roads conditions such as buoyancy is critical.	Available today in most European countries, but with different details.
Forest machine data	Current forest machines start to gather lots of data from machine condition, harvest and forwarder statistics. Other information, e.g. on trafficability may also be included and all together these datasources will play an important role to collect reliable and up-to-date information on forest operations.	Available today in some European countries. To be further studied in how to include in forest operations models or decision support.
VHR imagery (or UAV for more local analysis)	VHR imagery from satellite data have the potential to both be used as imagery and produce digital surface models and then for forest estimates. The coming VHR initiatives, possible with a new business model, may produce far more frequent data than current sensors and may therefore be a potential important data source.	Available today in all European countries, but not included in forest operations models or decision support.
Mobile laser scanning	Laser scanning produce reliable data in forest applications and more details may be captured with mobile laser scanners. Details may include both ground conditions as well as details from the stems.	Available today in most European countries, but only tested in R&D.

[32] Spatial resolution 1-5 meter


EFFORTE