Suomen ympäristökeskus
Finlands miljöcentral
Finnish Environment Institute

NATURAL RESOURCES
INSTITUTE FINLAND

Novel methods for the accounting of forest ecosystems and circular materials (2019-FI-ENVECO)

Methodological report / January 2021

European Statistics Action Grant (Environmental accounts and Ecosystem accounting)

Hurskainen Pekka, Salminen Jani, Balázs András, Johansson Annika, Kangas Annika, Karppinen Tiina K. M., Laturi Jani, Mäyrä Janne, Pohjola Johanna, Savolainen Hannu, Tuominen Sakari, Vihervaara Petteri, Virkkunen Henri

Contents

List of Tables

List of Figures

1 Introduction

The System of Environmental-Economic Accounting (SEEA) is an internationally agreed framework integrating economic and environmental data to provide a comprehensive and multipurpose view of the interrelationships between the economy and the environment and the stocks and changes in stocks of environmental assets, as they bring benefits to humanity. It brings together economic and environmental information in an internationally agreed set of standard concepts, definitions, classifications, accounting rules and tables to produce internationally comparable statistics (UN 2021).

The SEEA consists of two parts: the SEEA Central Framework (SEEA CF), which looks at “environmental assets”, such as water and energy resources, forests, raw materials, etc., their use in the economy and returns back to the environment in the form of waste, air and water emissions. One example account of the SEEA CF is material flow accounts, which provide information on material inputs and outputs of an economy, an important resource to those wanting a macro, economy-wide look at the physical inputs into an economy, material accumulation in the economy and outputs to other economies. The second part of SEEA consists of Experimental Ecosystem Accounting (SEEA EEA), which differs from the somewhat similar sounding environmental accounts in the SEEA CF by taking the perspective of ecosystems and the flow of ecosystem services to society, instead of looking only at natural resources. Ecosystem accounts enable the presentation of indicators of the level and value of “ecosystem services” in a given ecosystem accounting area (UN 2021). The SEEA EEA framework has gone through major revision and global consultation in 2020 and a draft version of the revised SEEA EA (Ecosystem Accounting without ‘Experimental’) guidelines was published (UNSD 2020).

Finnish Environment Institute (SYKE) and Natural Resources Institute Finland (Luke), the partners of the ENVECO consortium, are national statistical organizations which report to Statistics Finland for provisioning of various material, environmental and ecosystem accounts in Finland. The ENVECO project, co-financed by Eurostat action grants for Environmental accounts and Ecosystem accounting, was led by Finnish Environment Institute (SYKE) in partnership with Natural Resources Institute Finland (Luke) between January 2020 and January 2021.The first aim of the project was to develop novel remote sensing, machine learning and spatial analysis methods for accounting of forest-related ecosystems in the SEEA-EEA framework. As a result, usefulness and quality of source data for forest ecosystem service and ecosystem condition accounting was improved and some data gaps were filled by testing and development of forest variables and indicators. Another task related to forest accounts was to provide the value of forests by using monetary asset accounts. The value for timber production was calculated by using three approaches. The value of carbon sequestration was expressed as net present value for different prices of carbon.

The second aim of ENVECO was to address the data gap in existing material flow accounts which do not separate primary and secondary raw materials. Such distinction is essential to provide information on the volumes of the material recovery and recycling in the circular economy. As a result, the project contributed to the monitoring of the transition towards a circular economy in Finland, benefitting Statistics Finland in compilation of material flow accounts, decision makers and policy planners in the country, secondary material producing enterprises and the great public by quantifying and making the secondary material flows visible.

2 Circular economy (WP2)

Johansson Annika, Virkkunen Henri, Savolainen Hannu, Karppinen Tiina, Salminen Jani

Finnish Environment Institute (SYKE)

The work conducted in Work Package 2 (WP2) contributes directly to the monitoring of the transition towards a Circular Economy (CE) in the European Union. Material flow analyses were performed for a set of 14 materials, with the goal of identifying and quantifying material flows directed towards secondary material production. In addition, data gaps and inaccuracies were observed and highlighted. Based on the work, supply and use tables for the fourteen waste materials were created. This was done in a highly disaggregated manner which made it possible to observe the contribution of each specific industry to the material flow in addition to a broad economy-wide view. In addition, the tables are fully compatible with existing national material flow accounts.

The stated specific objectives in WP2 were:

- To address the gap in existing material flow accounts, that do not separate the primary and secondary raw materials. Such distinction is essential to provide information on the volumes of the material recovery and recycling in the CE. To measure the achievements related to this objective, the number of individual materials for which material flow accounts were generated were used as an indicator.
- To assess the availability, representativeness and quality of data for the secondary material accounts material by material. This way, the data gaps and future development areas could be recognized.
- To study the contribution of different industries to the processing and generation of secondary raw materials. For this objective, the level of sector disaggregation was provided as an indicator.

2.1 Background

Economy-wide material flow accounts (EW-MFA) practically exclude waste or secondary material flows (Eurostat, 2018). The focus of EW-MFA is to measure the material inputs into and material outputs of the economy, excluding material flows within the economy. Waste flows are measured only under two categories: waste traded for final treatment and disposal (physical imports and exports) and waste disposal (domestic processed output). This gives rise to develop measuring and accounting for material flows within economy, i.e., secondary and recycled material flows.

Recycling and recovery of materials is essential part of sustainable Circular Economy (CE). The key idea of CE is reducing the consumption of primary raw materials and promoting the reuse and recycling of materials already in circulation (secondary materials). In a CE system, waste disposal such as landfilling or incineration is considered a last-ditch measure that is only utilized if prevention of waste production, re-use or recycling is not feasible. The aim is to conserve/preserve the reserves of renewable and non-renewable natural resources, and to restrict the environmental impacts resulting from their uptake and use.

In order to facilitate the transition from a linear to a circular economy, the revised EU Waste Framework Directive (2008/98/EC) and its amendment (2018/851) set ambitious waste recycling targets. Regarding packaging waste, these include recycling targets of 85% for paper and cardboard, 80% for ferrous metals, 60% for aluminum, 75% for glass, 55% for plastics and 30% for wood. These goals are to be achieved by 2030–2035. For household waste, a minimum recycling goal of at least 55% by weight is set to be achieved by 2025. This goal becomes progressively more ambitious, increasing to 60% by 2030 and 65%

by 2035. The recycling target/rate goal for construction waste is 70%. Additionally, the new Waste Framework Directive also requires improvements in the collection of national waste data and waste accounts.

In order to monitor the transition to CE, disaggregated data on the generation of waste and its subsequent fate and use is required on a national scale. Such a disaggregation would be particularly valuable concerning waste with high volume and broad waste classes such as mixed construction and demolition waste (C&DW) or mixed municipal solid waste (MSW). For accounting purposes, it is necessary to be able to view the material flows both as a high-level overview spanning the economy but also retain the ability to “zoom in” to examine certain sectors of the economy. Waste accounting provides a methodological framework with which this aim can be achieved. Economy-wide material flow accounts (EW-MFA) practically exclude waste or secondary material flows (Eurostat, 2018). The focus of EW-MFA is to measure the material inputs into and material outputs from the economy, omitting material flows within the economy. Waste flows are measured only under two categories: waste traded for final treatment and disposal (physical imports and exports) and waste disposal (domestic processed output). This gives rise to developing measuring and accounting for material flows within economy, i.e., secondary material flows.

Despite the relatively high standard of waste material flow accounts currently compiled in Finland, unaddressed shortcomings include that the role and extent of secondary material supply and use cannot currently be statistically assessed.

To overcome this, we compiled and provided data for the supply and use of secondary materials. These were refined into secondary material accounts detailing material-specific supply and use, industry by industry. In addition, we examined potential data gaps and discrepancies that arose when the compiled data were compared to existing waste statistics. The outcomes of the developed method are compatible with current waste statistics collected and reported by Statistics Finland.

2.2 Methodology in WP2

2.2.1 Data collection

In order to achieve the goals outlined in chapter 1 we compiled data on the supply and use of wastes used for secondary material production. In order to improve the reliability of the data, several data sources were utilized in parallel. These included the following:

- Direct inquiries to ca. 70 companies, production units or organizations engaged in waste collection, processing and secondary material production. (Bottom-up data)
- Producer responsibility statistics (Pirkanmaa Centre for Economic Development, Transport and the Environment). (Top-down data)
- Data bank on the material composition of mixed municipal solid waste (hereafter referred to as mixed MSW), compiled by Suomen Kiertovoima KIVO ry, an umbrella organization for public waste management, treatment and disposal enterprises. (Bottom-up data)
- Sales data per enterprise, obtained via public database by Suomen Asiakastieto Oy. (Bottom-up data)
- Data on waste generation and handling reported in the VAHTI/YLVA compliance monitoring system managed by SYKE. Hereafter referred to as VAHTI/YLVA. (Bottom-up data)
- Amounts of imported and exported waste were obtained from the Finnish Customs Database. (Top-down data)

- Data on waste produced in building disassembly (construction and demolition waste, hereafter referred to as C&DW) were gathered from Building and Dwelling Register (BDR), maintained by Digital and Population Data Services Agency. (Bottom-up data)
- The above data was combined with on-site survey data on the material composition of demolished buildings provided by Ytekki Oy in order to calculate waste generation coefficients for different building types by construction material. (Bottom-up data)
- Data on material flows concerning fertilizer production were obtained from the Finnish Food Safety Authority. (Bottom-up data)
- Regarding slag from waste incineration, data previously collected by Statistics Finland from waste incineration companies were utilized. (Bottom-up data)
- Data on the use of sand, gravel, clay and crushed stone obtained from Confederation of Finnish Construction Industries RT. (Top-down data)
- Waste statistics collected and published by Official Statistics Finland.
- Data on MSW was obtained from "Municipal waste by treatment method in 1997 to 2017" (Official Statistics Finland). (Top-down data)

The data collection covered the chain extending from waste generation through waste collection and treatment to the production of secondary materials. Collection was conducted in a bottom-up, industry and material-specific manner, which could be additionally aggregated to describe the entire Finnish economy. Additional top-down data were also collected to allow comparisons between the supply and use tables. Industry subdivision is based a 26-category scheme utilized in the environmentally-extended input-output model ENVIMAT developed and managed by SYKE (Seppälä et al., 2011). This scheme was modified by adding Demolition of Buildings as a separate industry.

2.2.2 Data analysis and use

To facilitate the collection and analysis of data, we first defined the structure of the hierarchy concerning waste and secondary materials. The system included the following steps (Figure 1):

- Generation of waste by industry or source (primary waste)
- Waste leaving the economy as export or entering as import
- Waste collection and pretreatment
- Waste received and processed in secondary material production and reject thereof
- End-products: manufactured secondary materials supplied to the market

Existing data on wastes used in the production of secondary materials were collected and combined in order to create the accounts on the secondary material flows within the Finnish economy. The compiled industry-specific data could be used to examine how different industries contribute to waste generation and secondary material production. By observing data gaps and mismatches in the compiled data, sources of error and areas for further development were detected.

The data gathered from other sources were compared with the reported waste data in VAHTI/YLVA in order to further assess the reliability and robustness of the gathered data and to identify data gaps.

Supplementary data was utilized in order to better assess the contribution of certain sectors of the economy that the compliance system does not address sufficiently. These include sectors such as mixed C&DW, waste export/import and municipal solid waste (MSW).

The generation of MSW is not reported directly. Data on the treatment of MSW is retrieved from the national waste statistics and complemented with external data sources, for example for the material consistency of the mixed MSW (KIVO) and disaggregation between MSW from households and administration, service & business sector (Salmenperä et al. 2016).

VAHTI/YLVA contains information on the country of origin regarding treated waste. Likewise, similar information is provided on where the intended recipients for generated waste are located. However, since it contains only those operators with Regional State Administrative Agency permits and lacks all others, VAHTI/YLVA does not possess information on all shipments entering or leaving the country. For this reason, customs statistics were considered more comprehensive.

Like MSW, demolition data is indirectly reported in VAHTI/YLVA and reported in an aggregate with other waste from construction. For this reason, building demolition and renovation data from BRD and Ytekki Oy were utilized to assess this industry more specifically.

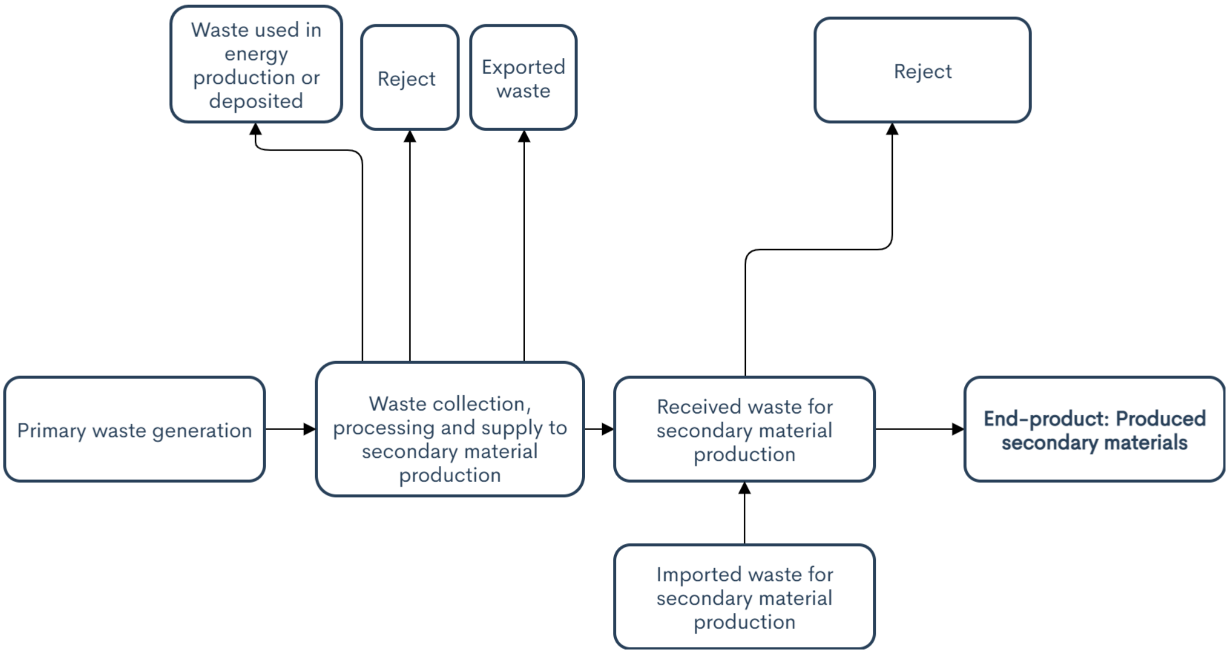

Figure 1: The studied steps of the waste-to-secondary material -cycle.

For this work package, a suite of 16 types of waste materials were analyzed.

- Glass
- Plastic
- Wood
- Iron and steel
- Aluminum
- Copper
- Paper
- Cardboard
- Concrete
- Bricks of clay
- Rubber
- Sand, gravel, clay and crushed stone
- Fertilizers
- Mineral fractions from slag
- Asphalt
- Sludge from municipal wastewater treatment

Each of these materials were found in different waste streams, such as industrial waste, C&DW and MSW. A data gap here lies in the fact that some categories (e.g. mixed MSW, mixed packaging materials and mixed C&DW) may in practice consist of several types of materials, the true compositions being unknown and cannot be thus included in the calculation without uncertainty. Research and databases were used to estimate the material consistencies of these mixed streams.

As explained above, for each stage of the waste-to-secondary material -cycle different and independent data sources were used. Thus, for each of these steps, independent figures on the volumes were generated. In principle, supply and use should always be equal. For instance, supply of specific waste to secondary material production should be equivalent to the use of that waste by secondary material manufacturers. In this report, however, we present unbalanced supply and use tables and discuss the reasons and methodological limitations related and contributing to this in the Discussion and conclusions section.

2.3 Material-specific summary accounts and findings on data availability and data gaps

In the following, we present summary supply and use accounts, findings on data availability and data gaps for each secondary material in separate sub-chapters. These results are also summarized in a separate document "Material flow accounts for individual secondary raw materials" which was a separate deliverable (D2.1) of work package 2. This deliverable also includes indicators describing data quality of each figure in the tables.

2.3.1 Concrete and bricks from clay

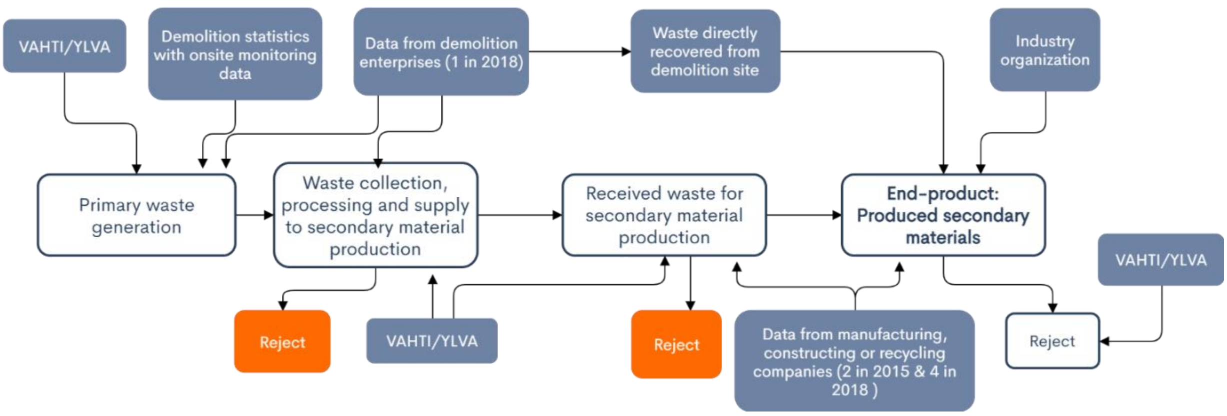

Figure 2: Data sources regarding concrete and bricks from clay in blue notes and orange notes indicating biggest data gaps/missing data. In parenthesis the number of companies from which information was obtained.

More data were available from the latter than the earlier phases of the concrete and brick waste-to-secondary material -cycle. Regarding *primary waste generation*, data on generated concrete and brick waste were collected via direct inquiries from demolition enterprises and manufacturing, construction and recycling companies (Figure 2). Additional data on concrete and brick waste resulting from demolition of buildings were produced by combining figures on demolished buildings in 2015 and 2018 as reported in the register of buildings and dwellings (BDR) with an on-line monitoring dataset on materials generated from demolished buildings of different use types and characteristics (e.g. main construction materials).

Uncertainties in the primary waste generation data related to the missing and erratic data on the demolished buildings in BDR. Most often the missing parameters were gross floor area and/or construction materials. Extent of the on-site monitoring data on the material composition of the demolition waste was limited and thus insufficient to account for all material combinations present in the BDR data. The on-site data accounted for material combinations of wood and concrete structures, with other material categories (steel, unknown, other) being absent from the data. Classes such as “Unknown” and “Other” are prevalent in BDR data and cannot be used to provide material-specific assessments. Generation of more comprehensive data sets on the actual material composition of demolished buildings would improve the accuracy of these estimates.

Besides this, another potential source of uncertainty regarding waste generation was the relatively high amount of buildings (typically small private storages, garages, detached saunas and other miscellaneous buildings) with unknown construction materials and/or unreported gross floor areas. Even if waste amounts from an individual building such as these may be small, their prevalence in the data means their combined contribution to waste generation may be significant. Furthermore, there were also uncertainty concerning the amount of brick and concrete present in mixed C&DW. Due to the limited number of responses to the company surveys and the lack of survey data, the assessment could not be performed with sufficient reliability. To conclude, first, our primary waste generation data underestimate the actual volume of concrete and brick waste volumes. Second, more research data is needed to generate more reliable accounts with better coverage for the generation of concrete and brick waste.

Major data gaps were evident regarding the supply-side of brick and concrete waste, with practically no direct data available on primary waste generation. Based on indirect, BDR-derived building demolition data, we estimated that circa 496 750 tons of concrete waste and 55 195 tons of brick waste were generated in demolition in 2015. In 2018, 614 540 tons of concrete waste and 68 282 tons of brick waste were generated, respectively. Based on the information provided in VAHTI/YLVA, construction and demolition were the highest (in year 2015) and third-highest contributing (in year 2018) industry to the generation of concrete and brick waste. In the latter case, highest and second highest contributors were “Waste treatment and disposal” and “Manufacture of articles of concrete, cement and plaster”, respectively. This indicates that despite the lack of other data, the supplementary information could be utilized in its place to reduce the extent of the data gap on primary waste generation.

Regarding *waste collection, processing* and *supply to secondary material production* from construction, concrete and brick waste is either transported to off-site processing or some of it gets treated directly at the construction site. The permanent facilities processing concrete waste receive concrete waste not only from demolition but also from the manufacturing of products of concrete and ready-mixed concrete. In both cases, the concrete waste is mainly recovered by using the crushed material in earth works (Nylén and Salminen 2019). The data on the amount of crushed concreted i.e. *the amount of produced secondary material*, were obtained from individual enterprises and calculated by using the company’s own estimate of their market shares. The data cover both on-site and off-site processing and cover in total roughly 25% in 2015 and 53% in 2018 of the market as based on the summed-up market share estimates made by the companies. The responses to the business surveys were satisfactory and the total volumes estimated from company-specific data showed only moderate variance (1 605 911 – 1 777 615 t/a). However, a higher response rate would have improved the coverage and reliability of the data.

Finnish national legislation[1] permits concrete and brick waste to be used in earth works by using a registration protocol replacing the need for an environmental permit (Nylén and Salminen 2019). These volumes of recovered concrete and brick waste must be reported to the state authority case by case. However, the data were not readily available as the electronic system for the reporting the recovered volumes was

[1] Regulation on the utilization of certain wastes in earth construction (843/2017), repealing regulation (591/2006)

introduced in 2020. In Pirkanmaa, Pohjois-Pohjanmaa, Uusimaa and Varsinais-Suomi regions, which together cover roughly 55% of the country's population, in total 2,0 million tons of concrete waste was recovered by using the above protocol in 2018. These figures cover crushed concrete from both on-site and permanent (off-site) treatment facilities. They, however, cannot be used as such to calculate the national figures as Pirkanmaa, Pohjois-Pohjanmaa, Uusimaa and Varsinais-Suomi regions have much higher construction activity in comparison with the average of the rest of the country. Also, some of the registered cases may have not been realized and the volumes actually recovered may be differ (usually by being lower) from that given in the registration form. That said, the figures may overestimate the total annual volume of concrete waste recovered. From 2020 onwards, data availability on the recovery of the wastes belonging to the scope of the governments decree (including crushed concrete and bricks) for earth works will enable more comprehensive accounting on these materials in Finland. Accessing the data on the amount of waste treated on-site proved most challenging within this domain as data were not systematically summarized by companies. However, data availability was somewhat better for 2018 allowing an estimation based on data obtained from a company and scaled to cover the industry by using revenue information. For 2015 such estimate could not be generated as data for that year were completely missing.

To account for *reject* from secondary material manufacturing, the amount of non-recyclable concrete and brick waste deposited to landfill were obtained from VAHTI/YLVA. There were no data readily available to estimate the amount of waste rejected at other stages of the waste-to-secondary material cycle. However, these wastes can in principal only be deposited to landfills that have the obligation to report to VAHTI/YLVA. Hence, the data coverage should be good.

Concrete and brick waste in VAHTI/YLVA were sometimes reported together as an aggregate class. The proportion of mixture wastes varied greatly between years. In terms of tonnage, less than 5% of all generated concrete and brick waste in VAHTI (2015) were mixtures, while in YLVA (2018) the corresponding percentage was significantly higher at 26%. These classes were, however, still utilized relatively often: 19% of all encountered concrete waste classifications were mixtures in 2015, and 30% in 2018. There were very few instances of other waste reported erroneously as concrete and bricks, which means that at a basic level the data can be considered reliable. However, concrete, bricks and their mixtures are currently reported as a summed aggregate class in the system. For an accurate material flow account, these needed to be manually reclassified in order to separate the three material flow components, which was difficult and labor-intensive.

For future development, reporting material flows of concrete, bricks and mixtures as separate entities instead of an aggregate class or the use of machine learning or other methods to distinguish between the different materials in data systems would make it less work-intensive and more accurate to compile economy-wide waste statistics. Due to the nature of demolition practice, it may not be possible to keep the materials separate and thus further research on the most common brick-to-concrete ratios in the mix could improve the accuracy of data used in waste statistics.

Based on information received from companies, they were present in their mixtures in proportions 90% and 10%, respectively. This was further validated with information from VAHTI/YLVA, based on which 10% was used as the share of brick in the waste mixtures from demolition and 6% from other construction work. Aggregate classes are prevalent in the data as they are sufficient for VAHTI/YLVA's original role in environmental permit compliance monitoring. However, for accounting purposes they present a source of inaccuracy.

Table 1: Merged summary use and supply table for concrete and brick waste in 2015 and 2018 in Finland.

Material	Use in secondary material manufacturing*	Reject from secondary material manufacturing, t	Supply of secondary material, t
Concrete (2015)	1 633 524	23 274	1 708 431
Concrete (2018)	1 421 774	80 146	1 652 842
Bricks from clay (2015)	104 268	1 795	109 049
Bricks from clay (2018)	157 974	2 178	183 649

* Use in secondary material manufacturing includes waste received for industrial manufacturing of secondary materials and waste treated directly at demolition sites (on-site treatment).

Regarding material flows in export and import phases of waste hierarchy, concrete and brick waste were neither exported to nor imported from or to Finland in 2015 or 2018 and imports and exports are therefore omitted from Table 1. In addition, concrete or bricks were not supplied to waste-to-energy plants as these are not suitable for that purpose. The amount of reject includes the amount of contaminated concrete and brick waste obtained from VAHTI/YLVA.

Regarding waste received for secondary material manufacturing (i) and produced secondary material (end-product) (ii), the amounts of (i) concrete and brick waste supplied to secondary material manufacturing and (ii) secondary materials produced were higher when calculated on the basis of company inquiries. These figures were also considered more accurate than those derived from VAHTI/YLVA. The reasons for this relate to – as mentioned above – the absence of companies treating concrete and brick on-site are not reporting earth works in the system. The amount of brick waste was higher in 2018 as the company data in 2015 were mainly obtained from permanent treatment facilities and thus 6% instead of 10% was used for the ratio of brick waste. The difference between raw materials for secondary material manufacturing and end products produced were due to changes in storage.

2.3.2 Paper and cardboard

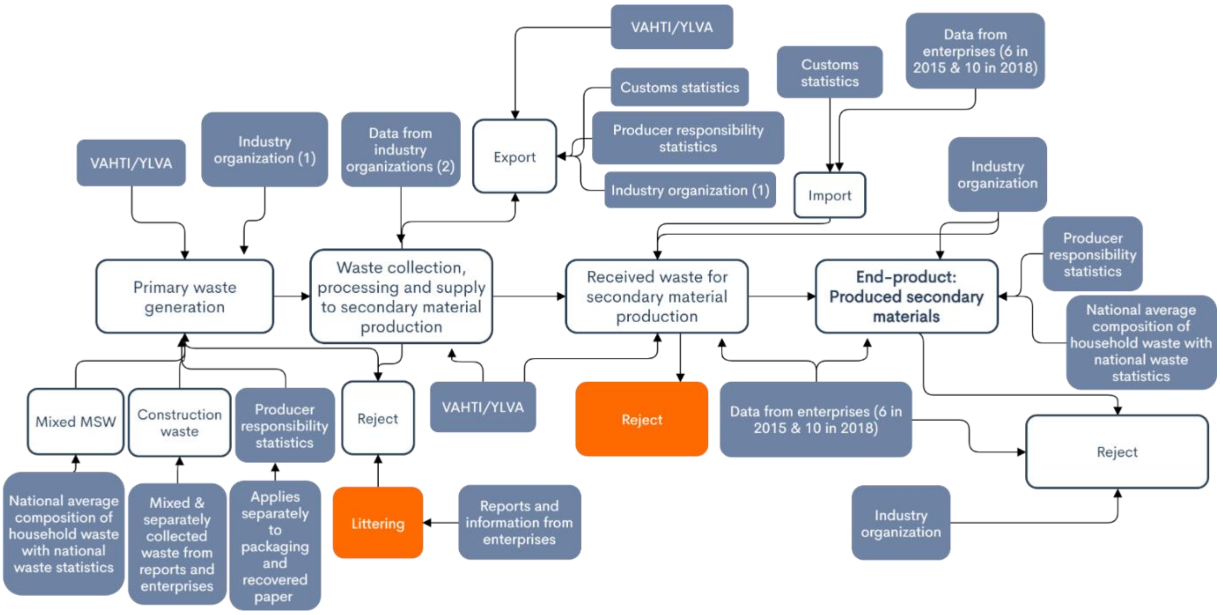

Figure 3: Data sources regarding paper and cardboard in blue notes and orange notes indicating biggest data gaps/missing data. In parenthesis the number of companies from which information was obtained.

Comprehensive data on paper and cardboard waste streams were obtained, especially towards the last phase of the material cycle, as recycling companies responded comprehensively to the inquiries and the response rates were high (Figure 3). Companies were well aware of the amounts of waste volumes they received, as well as of the rejects and the amounts of recycled materials produced. Data on the volumes of waste received and secondary materials produced were additionally obtained from two industry organizations involved e.g. in arranging the producer responsibility and the figures corresponded to the data obtained from VAHTI/YLVA and reported by the individual companies.

Data gaps were mainly identified in the assessment of the amount of *primary waste generated* and on *reject from generation and collection* stages of the waste-to-secondary material -cycle. For example, data on generated municipal waste were not available because waste producers (residents, majority of companies, public sector, etc.) do not report waste volumes in the VAHTI/YLVA database, and therefore the assessment was based on data from waste reception facilities and waste statistics. Also, data on the packaging material flow resulting from online retail entering the market were unavailable. For future development concerning the generation of packaging waste, the volumes of packaging material provided for online retail trade could possibly be obtained from packaging material wholesale enterprises. A supplementary source of information could be the largest postal and courier services responsible for transporting the product for end use. However, this would not consider the retail packaging of the product itself, only the packaging used for transport.

Separately collected paper and cardboard waste generated during construction is included in the producer responsibility data, which concerns plain paper and paper and cardboard packaging separately. Paper and cardboard in the mixed MSW fractions and C&DW were expected to end up in incineration (although before 2016, mixed waste also ended up in landfills) and as such was allocated in the column 'energy production' in Figure 3. Their shares were calculated based on national estimates on MSW composition.

The amount of paper and cardboard waste generation in litter and in use in e.g. small-scale household combustion were impossible to determine, because there were no data available to identify the latter, and although information was available on the amount of street sweeping waste from the companies carrying out the activities, its composition remained unknown. However, the contribution of these flows to the overall flow of paper and cardboard in the Finnish economy is likely small.

Regarding *import and export* of waste (Table 2 and Table 3), the data for exported waste were obtained from customs statistics. As the exported amount of paper and cardboard waste were reported together, the total amounts of each component were estimated based on the paper/cardboard-ratios obtained from the producer organization statistics (59% Paper and 41% cardboard in 2015, 50% paper and 50% cardboard in 2018, respectively). The data for exported waste were in line with the data from VAHTI/YLVA for 2015. The data for imported waste were obtained from companies and were in line with the customs statistics for both years in question. Data gaps in this phase of the waste-to-secondary material -cycle are that small waste shipments inside EU do not require a customs declaration. If shipments such as these also have no waste shipment permits or obligation to report waste exports in the environmental permit, there is no information available on such shipments.

Data on the amount of *waste received by companies for production of secondary materials* (Table 3) were obtained from VAHTI/YLVA, the industry organization and the companies concerned. These sources of information were determined to be consistent with each other. Paper and cardboard were reported as an aggregate group in VAHTI/YLVA. This data contained large amounts of mixed waste, both in terms of occurrences and tonnages. In 2015, circa 36% (VAHTI) and in 2018 circa 40% (YLVA) of all generated paper and cardboard waste tonnages were mixtures. In both data sets, mixture-type classifications were common, accounting for 29% of all encountered classifications in 2015 and 26% in 2018. Because of this, a manual recoding was necessary to enable disaggregation similarly to the case of concrete and brick waste.

A higher-resolution classification, in which paper, cardboard and mixtures are separately reported, would enable higher accuracy of the waste and secondary material accounts. However, it would be challenging to implement as for instance paper and cardboard packaging waste is collected from households in mixed form. The issue of aggregated data was not present in the data obtained by direct inquiries, since the enterprises mostly dealt with either paper or cardboard and only rarely both. As such, the inquiry data collected did not require recoding.

Concerning cardboard *waste use for secondary material production* in 2018, data were comprehensively obtained from companies using cardboard waste and with almost similar coverage regarding paper waste utilizing facilities in 2018 and 2015. The company inquiries did not provide as comprehensive data regarding cardboard waste in 2015 and as such the accuracy of the 2018 data could be verified more reliably than data for 2015. Regarding cardboard waste in 2015, data were obtained from VAHTI/YLVA and compared to data from industry organization together with company data, because data were not available from many companies due to changes in company's data storage methods.

Regarding *reject* from paper and cardboard secondary material manufacturing, the following estimates based on direct company inquiries were utilized. In 2015 and 2018, paper waste contained 1-31% of reject. Based on this 20% was used as the share of reject for data obtained from VAHTI/YLVA in 2015 and 2018. On the basis of company responses concerning year 2018, 10% was used as the share of reject from cardboard waste in 2015 and 2018. The data on *the production volume of secondary materials* from paper and cardboard waste in 2015 were equal to data from national statistics. The reject consists misclassified waste materials such as plastic or metals for example, i.e. newspaper staples. The rest of the waste was utilized in the production of secondary materials, for example in newsprint, case materials, household and sanitary paper.

A higher-resolution classification, in which paper, cardboard and mixtures are separately reported, would enable higher accuracy in creating waste accounts. However, it would be challenging to implement as for instance paper and cardboard packaging waste is collected from households in mixed form. The issue of aggregated data was not present in the data obtained by direct inquiries, since the enterprises mostly dealt with either paper or cardboard and only rarely both. As such, the data collected in the inquiries did not require recoding.

Table 2: Merged summary supply table for paper and cardboard waste in 2015 and 2018 in Finland.

Material	Domestic supply, t*	Import, t
Paper (2015)	367 187	37 130
Paper (2018)	277 803	82 586
Cardboard (2015)	265 121	23 346
Cardboard (2018)	273 833	25 500

* Collected paper and cardboard waste

Table 3: Supply of paper and cardboard waste to energy production and use of paper and cardboard in waste secondary material manufacturing, reject in secondary material manufacturing and the supply of secondary paper and cardboard in 2015 and 2018 in Finland.

Material	Export, t	Energy production, t	Waste received to secondary material manufacturing, t	Reject in secondary material manufacturing, t	Produced secondary material, t
Paper (2015)	54 392	89 964	312 075	69 527	242 548
Paper (2018)	34 335	89 392	313 717	69 685	244 033
Cardboard (2015)	37 798	116 597	269 786	26 979	242 807
Cardboard (2018)	34 335	132 767	322 511	64 703	257 819

2.3.3 Plastics

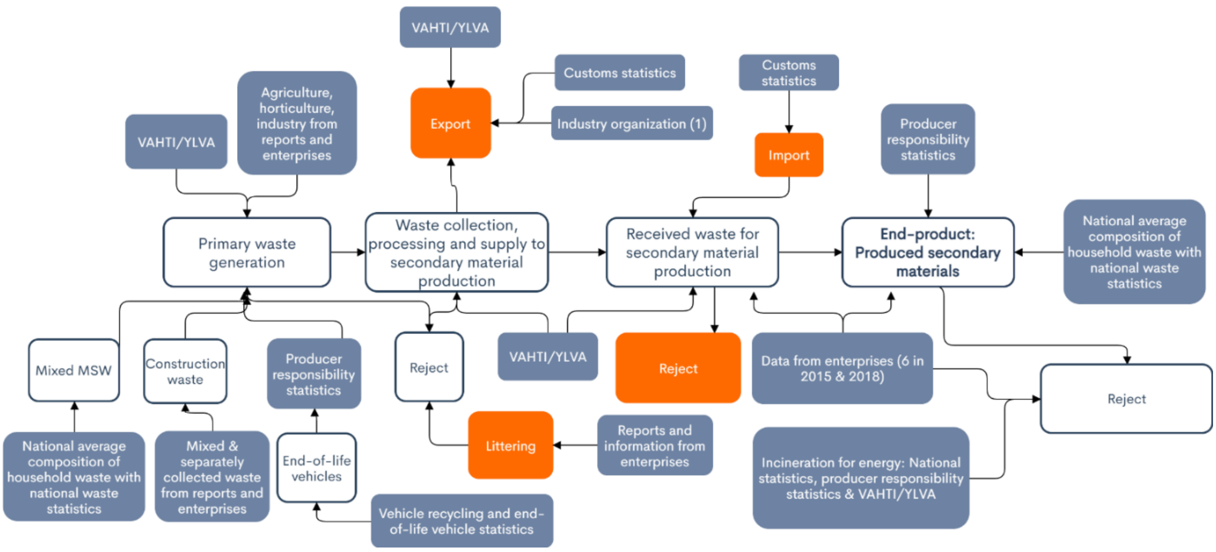

Figure 4: Data sources regarding plastic in blue notes and orange notes indicating biggest data gaps/missing data. In parenthesis the number of companies from which information was obtained.

The separate collection of plastic packaging waste from households in Finland is a relatively new practice that has only started properly after 2016. Prior to this, majority of plastic packaging has ended up in the mixed MSW. Plastic waste therefore ends up in secondary material production in relatively small quantities. Significant amounts of plastic packaging, as well as other municipal plastic waste, still ends up in MSW fractions which are generally incinerated (and additionally, before 2016 were also landfilled). As a result, there were gaps evident in the data concerning various stages of the waste hierarchy (Figure 4). More comprehensive data collection could enhance the efficiency of plastic recycling and help direct higher material volumes towards secondary material cycle.

The amount of *generated plastic waste* could not be estimated with confidence due to existing severe gaps in data availability, and therefore, the assessment was based on data from waste reception facilities and describes the waste collection stage of the material cycle.

The problem stems from missing pieces of data on several waste flows. For example, data on generated plastic waste included in MSW were not available, because waste producers (e.g. households and majority of private companies and public sector organizations) do not report their waste volumes into the VAHTI/YLVA database.

It was also impossible to determine the amount of generated plastic waste in litter because, for example, the composition of sweeping and roadside waste is unknown. Quantitative data on littering were collected from individual companies, but an inventory of the amount and composition of littering is needed in order to be able to take this material flow into account. However, the new reporting requirements from the EU directive on single-use plastics (2019/904/EU), leads to increased monitoring of gardening, sweeping and roadside waste.

Another data gap in plastic waste generation concerns the generation of plastic waste from agriculture. About 12 000 tons of agricultural plastic waste is generated per year, part of which is in producer responsibility statistics. It has been estimated that about one-fifth goes to recycling and the rest is used for energy. (Alenius 2016) In 2015-2019 there were no functioning system for the recycling of agricultural plastic waste, and recycling was largely based on the actions of individual farmers and horticultural enterprises (Erälinna & Järvenpää 2019).

Some data gaps were still evident concerning *waste collection*, even though producer responsibility statistics do provide rather comprehensive data on plastic packaging waste collected in Finland. Data were missing for instance, from companies with a less than one-million-euro turnover and data on the packaging material flow resulting from online retail entering the market. As with paper and cardboard, packaging wholesale enterprises and postal & courier services could in future provide sources of information for accessing this material flow. However, this approach would only consider the packaging used in transporting the product after sale, while the retail packaging of the product itself would remain unknown. Due to the immense variety of possible retail packaging compositions and scattered distribution of suppliers, it is likely that this material flow will remain unelucidated for the foreseeable future.

Plastic waste has been estimated to be the second largest material fraction of mixed household solid waste (second to biowaste) (KIVO) and a large fraction of mixed C&DW, together with wood, crushed stone and other miscellaneous waste (Liikanen et al. 2018). The assessment of the composition of mixed waste was based on national waste statistics, household waste composition database (KIVO), reports and company surveys and more research are needed to enable even more accurate assessment of its composition. Also, an increased efficiency in MSW and C&DW sorting may provide a method with which to redirect more of this waste flow to the secondary material cycle. Where information on the composition of household waste was available, information on the composition of mixed MSW from other sources, such as administration, services and businesses, most likely varies greatly and is currently largely unknown.

The *import and export* data were obtained from the customs statistics. Uncertainty regarding these data relates to unregistered small waste shipments inside EU, that do not require a customs declaration, a waste shipment permit or have the obligation to report waste exports in the environmental permit. The new amendments to the Basel Convention regarding plastic waste came into force in early 2021 and will improve this, as more waste shipments will be subject to a permit. More comprehensive statistics on plastic waste imports and exports will improve the reliability of these waste statistics.

Table 4: Merged summary supply table for plastic waste in 2015 and 2018 in Finland.

Material	Domestic supply, t*	Import, t
Plastic (2015)	41 791	1 317
Plastic (2018)	42 008	3 859

*Collected plastic waste, based on Producer responsibility statistics

Domestic supply of separately collected plastics waste (Table 4) was estimated at 41 791 tons and 42 008 tons, for 2015 and 2018 respectively. Of these, the amount of imported waste were 1 317 and 3 859 tons, respectively (3% and 8,4%). Domestic supply refers to Statistics Finland's data regarding separately collected municipal plastic waste. The data lacks other plastic waste collected separately from, for example, agriculture and construction, which could also end up being recycled. Plastic waste among mixed waste is assumed to end up in incineration, i.e. energy production, even thou in 2015 some of it might have also ended up in landfill. The data below regarding plastic used in energy production, refers to the estimated amount of plastic among the mixed MSW and mixed C&DW. This data lacks, for example, the incineration of plastic waste from agriculture and separately collected plastics.

Table 5: Supply of plastic waste to energy production and to secondary material manufacturing, reject in secondary material manufacturing and the supply of secondary plastics in 2015 and 2018 in Finland.

Material	Export, t	Energy production, t	Raw materials for secondary material manufacturing, t*	Reject in secondary material manufacturing, t	Produced secondary material, t
Plastic (2015)	31 277	270 816	15 304	1 189	14 268
Plastic (2018)	21 542	291 980	33 997	6 753	24 541

*waste received for industrial manufacturing of secondary materials

As the recovery percentages of plastic waste remain relatively small, this were also reflected in the relatively small absolute tonnages of secondary plastic materials manufactured (Table 5). Recycling rate in 2018 was almost twice that of 2015, which reflects the more efficient collection of separately collected plastic packaging from households since 2016. The recycled amount of plastic waste in 2015 corresponds to the data provided by the producer responsibility organization, while in 2018 the organization's data covered only about half of the estimated amount. Producer responsibility statistics give double and three times the amount of recycled materials, in 2015 and 2018 respectively. The detailed data from companies manufacturing secondary plastics (granulates or flakes) are expected to give a more reliable estimate of the actual amount of recycled material. The figures regarding raw materials for secondary material manufacturing, reject and produced secondary material are based on data collected from identified six major companies producing secondary materials from plastic waste. However, there are approximately 530 companies manufacturing plastic products in Finland according to the Finnish Patent and Registration office (PRH) and if some of these companies were utilizing plastic waste in their products in 2015 and 2018, the figures cannot be considered comprehensive. More detailed information is needed on the amount of plastic waste produced and used by companies.

VAHTI/YLVA database categorizes plastic waste into only two categories, plastic packaging and aggregated "other plastic wastes". Currently, a more accurate plastic waste accounting considering various plastic types and their potentially different end-use scenarios is not possible with these tools. A more precise definition of the data in the database would allow it to be better utilized in waste statistics.

2.3.4 Metals

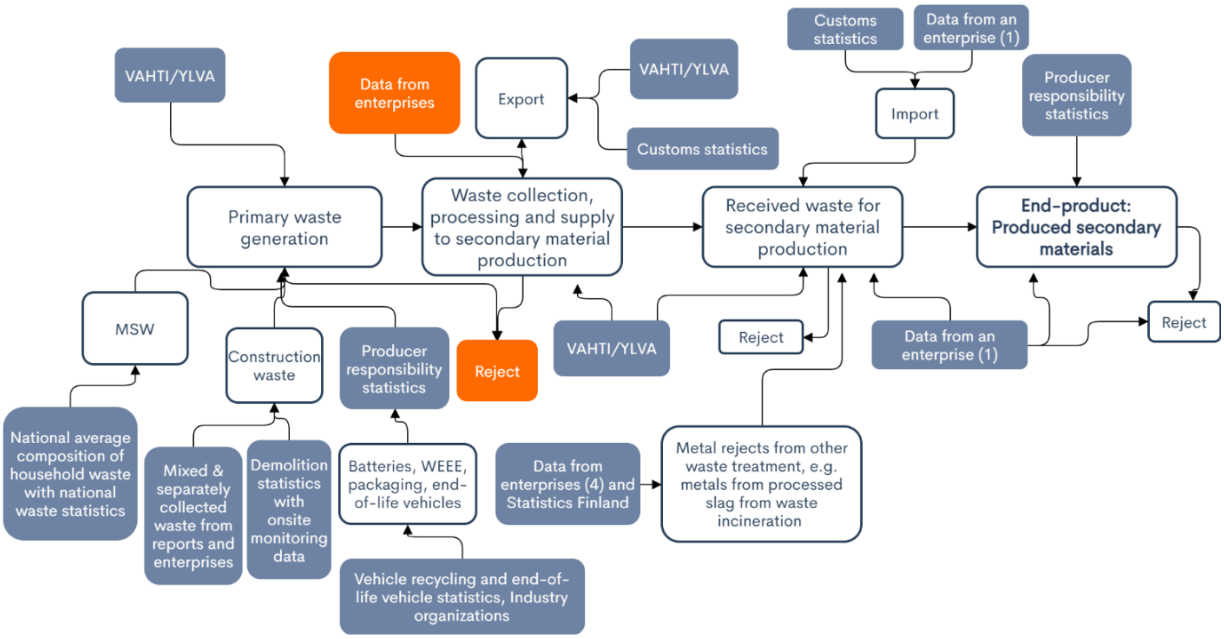

Figure 5: Data sources regarding metals in blue notes and orange notes indicating biggest data gaps/missing data. In parenthesis the number of companies from which information was obtained.

Data on the amount of *received metal waste for secondary material production* were considered comprehensive, unlike data on the beginning of the waste-to-secondary material -cycle, i.e. generation, collection and processing before the supply to secondary material production and reject from these activities were lacking some available data (Figure 5).

The amount of metal *waste generated* could not be estimated with confidence due to existing severe gaps in data availability. Like with many waste materials, this could have been done by using data from later stages in the waste chain, but this project sought to find primary sources of information. Uncertainties regarding demolition statistics related to the absence or generalizations of recorded building properties and on-site monitoring covering only a limited number of building types. Generation of more comprehensive data sets on the actual composition of demolished buildings would improve the accuracy of these estimates.

Metal waste has been estimated to cover approx. 2% of mixed household solid waste and 4-12% of mixed construction and demolition waste from which 8% was used in this study (KIVO, Liikanen et al. 2018). The assessment of the composition of mixed waste was based on national waste statistics (Statistics Finland), database on the composition of mixed MSW from households (KIVO), reports and company surveys. More research is needed to enable more accurate assessment of its composition. The metal contained in the mixed waste is collected for recycling either before or after incineration.

The metal waste included in producer responsibility schemes, such as batteries, WEEE (Waste electrical and electronic equipment), packaging and end-of-life vehicles, were estimated using producer responsibility statistics for the amount of certain wastes and material estimates for cars were based on Suomen Autokierrätys data, WEEE on El-Kretsen AB 2019 data and batteries on Suomen Akkukierrätys 2017 data. Information on metal packaging waste was obtained from the producer organization (Mepak-kierrätys Oy). The uncertainty of the data results from the fact that not all municipalities and private companies report their data on metal waste they collect to Mepak. Small metal objects are not subject to producer responsibility and information on their quantities were not readily available.

Regarding *waste collection, processing and supply to secondary material processing*, data were obtained from VAHTI/YLVA (Table 6). Unfortunately, no information was available on the quantities of metal received, processed and forwarded by scrap dealers. Obtaining this information would have made it possible to compile comprehensive statistics on the collection of metals and particularly give data in regards of reject during collection and processing.

The *import and export* volumes of metals were based on customs statistics. These were manually recoded into broader classes to separate metal types iron and steel, copper, aluminum, and mixed metals (Table 6). The import and export volumes of metals are affected by fluctuations in the monetary value of metals worldwide.

The amount of *received waste for secondary material production* (Table 7) was obtained from VAHTI/YLVA and the amount of *reject* from the company survey together with VAHTI/YLVA database. The reject consists of (i) losses in recovery processing and (ii) metals that were contaminated e.g. with radioactive substances.

According to our data and the nature of the material, metal waste was not utilized in energy production.

Table 6: Merged supply table for metal waste in 2015 and 2018 in Finland.

Material	Domestic supply, t	Import, t
Metals (2015)	466 613	851 546
Metals (2018)	690 146	910 722

* collected metal waste

Table 7: Supply of metal waste to secondary material manufacturing, reject in secondary material manufacturing and the supply of secondary metals in 2015 and 2018 in Finland..

Material	Export, t	Domestic waste delivered to secondary material manufacturing, t*	Waste received to secondary material manufacturing, t**	Reject in secondary material manufacturing, t**	Produced secondary material, t**
Metals (2015)	451 812	155 688	1 366 189	26 418	1 339 821
Metals (2018)	622 536	269 262	1 471 347	24 200	1 383 763

*Based on VAHTI/YLVA data. ** Secondary material manufacturing took place in basic metal manufacturing (NACE 241-245)

Table 8: Import and export of metal waste and receival of metal waste to secondary material manufacturing in 2015 and 2018 in Finland.

Material	Import, t	Export, t	Waste received to secondary material production (incl. import), t
Iron & Steel (2015)	813 728	358 207	1 295 571
Aluminium (2015)	13 244	52 123	79
Copper (2015)	19 878	22 807	38 267
Mixed metals (2015)	1 373	181	32 322
Iron & Steel (2018)	877 494	489 728	1 391 792
Aluminium (2018)	12 803	80 139	30
Copper (2018)	11 020	29 676	41 376
Mixed metals (2018)	2 837	4 472	42 662

The amounts of iron and steel, aluminum and copper waste *imported and exported* were obtained from customs statistics and the amount of received waste for secondary material production were obtained from VAHTI/YLVA together with data on the reported primary industry of the enterprises (Table 8). In some cases, VAHTI/YLVA categorizes metallic wastes into too broad classes to be immediately useful for accounting purposes. Aggregated, mixed classes such as *Mixed metallic packaging* and *Other mixed metallic wastes* were difficult to use, since only their total amount was readily available and manual disaggregation was necessary. Additionally, metal components such as aluminum may be included with waste that was categorized with other waste classes. This makes it more difficult to monitor the recycling goals relating to these kinds of metallic waste.

2.3.5 Asphalt

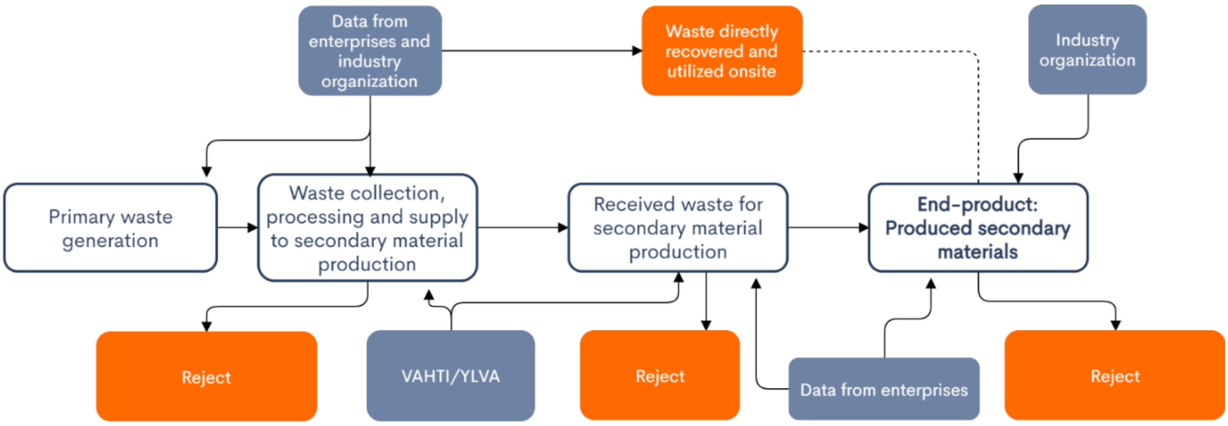

Figure 6: Data sources regarding asphalt in blue notes and orange notes indicating biggest data gaps/missing data.

The asphalt waste generated refers to waste asphalt that has been transported to an asphalt plant for recycling. However, for the operation of asphalt plants a registration protocol is used instead of an environmental permit. The registration covers the use of asphalt waste as well and therefore, the data on asphalt plants and the related use of asphalt waste is not present in VAHTI/YLVA. Asphalt waste that is immediately remixed on site as part of repaving, that is, making a new asphalt layer, is not considered waste. These on-site remix amounts (roughly 1 million tons per year; Forstén 2015) are therefore not considered as secondary material production.

The data on the amount of generated waste and produced secondary materials were obtained from enterprises and industry organization (Figure 6). The uncertainty was not related to the amount recycled but to whether the recycled asphalt could be considered waste before it was recycled or not. The amount in VAHTI/YLVA) were much lower for the reasons mentioned above and the information retrieved from the other sources were considered more reliable. Regarding the supply of asphalt, no imported waste was included in the data and all supply of asphalt waste was of domestic origin (Table 9).

Table 9: Merged supply table for asphalt waste in 2015 and 2018 in Finland.

Material	Domestic supply, t
Asphalt (2015)	1 800 000
Asphalt (2018)	1 900 000

Table 10: The use and production of secondary asphalt from asphalt waste in 2015 and 2018 in Finland.

Material	Produced secondary asphalt, t
Asphalt (2015)	1 650 000
Asphalt (2018)	1 900 000

For the use and production table, only the amount of final product was available, with 1.65 Mt and 1.9 Mt of secondary asphalt product produced in 2015 and 2017, respectively (Table 10). To our knowledge, asphalt is neither exported nor used in energy production. Data on the amount of asphalt waste used in secondary material manufacturing, or reject thereof, was unavailable.

2.3.6 Rubber

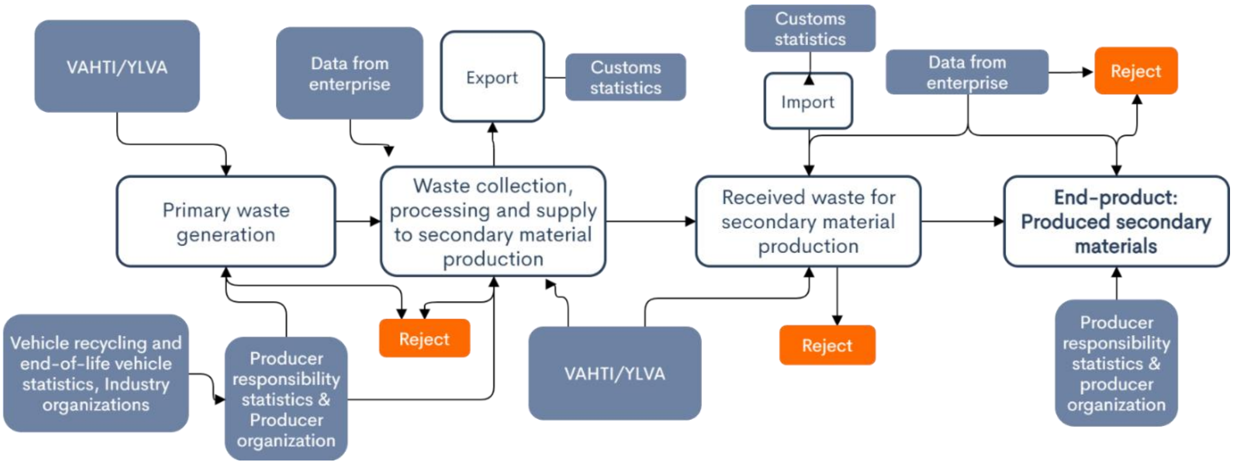

Figure 7: Data sources regarding rubber in blue notes and orange notes indicating biggest data gaps/missing data.

The amount of rubber *waste generated* could not be reliably estimated even though rubber waste generated by manufacturing industries was retrieved from VAHTI/YLVA (Figure 7). However, data on vehicle tyre waste generation is not reported to this database as a majority of it is generated by households and companies or organizations that do not have an environmental permit. Instead, data were available at the stage of *waste collection and processing* from producer responsibility statistics, producer organization (Suomen Rengaskierrätys Oy) and from a company that has been solely responsible for the recycling of rubber waste under producer responsibility in Finland. The *domestic supply* (Table 11) refers to data on producer responsibility statistics regarding collected tyres (Pirkanmaa Center for Economic Development, Transport and the Environment, Producer responsibility statistics, statistics on vehicle tyres). The import and export volumes of rubber for 2015 and 2018 are based on customs statistics.

For rubber, most of the total supply was accounted for by domestic supply, with imported material receiving only small shares in 2015 and 2018, 12% and 6%, respectively.

Table 11: The summary supply table for rubber waste in 2015 and 2018 in Finland.

Material	Domestic supply, t*	Import, t
Rubber (2015)	58 924	7 776
Rubber (2018)	60 808	3 446

* Share of collected used tyres: 55 453 t in 2015 and 57 152 t in 2018. The remaining 3 471 t (2015) and 3 656 t (2018) represent other rubber waste generation from manufacturing industries (VAHTI/YLVA). Of this waste,1226 t was utilized in secondary material manufacturing in 2015, and 1196 t in 2018.

The data *on raw materials for secondary material manufacturing* (Table 12) were obtained from producer responsibility statistics on collected tyres and from producer organization on energy production. There were no data available to estimate the amount of reject during waste collection, processing, supply to secondary material production and in receiving and producing secondary materials.

Uncertainties related to other rubber waste not covered by producer responsibility, such as rubberized rollers and retreading of heavy vehicles tyres and including these in the producer responsibility could be a solution in improving the reliability of the accounts.

Table 12: Export, energy recovery, re-use, material and other recovery of rubber waste in 2015 and 2018 in Finland.

Material	Export, t	Energy production, t	Re-use of vehicle tyres, t*	Material recovery, t*	Other recovery, t*
Rubber (2015)	1 841	13 571	831	34 567	6 459
Rubber (2018)	1 676	9 443	642	57 386	2 030

* Contains used tyres only.

In VAHTI/YLVA database, subclass of waste (used tyres) was the only one that is guaranteed to characterize the generation of rubber waste. Rubber that may be included with other classes cannot be separated without manual disaggregation. Data on rubber reject from secondary material manufacturing was unavailable. The data suggest that the share of rubber utilized in energy production has decreased over time between 2015 and 2018, with more being directed to secondary material production.

2.3.7 Wood

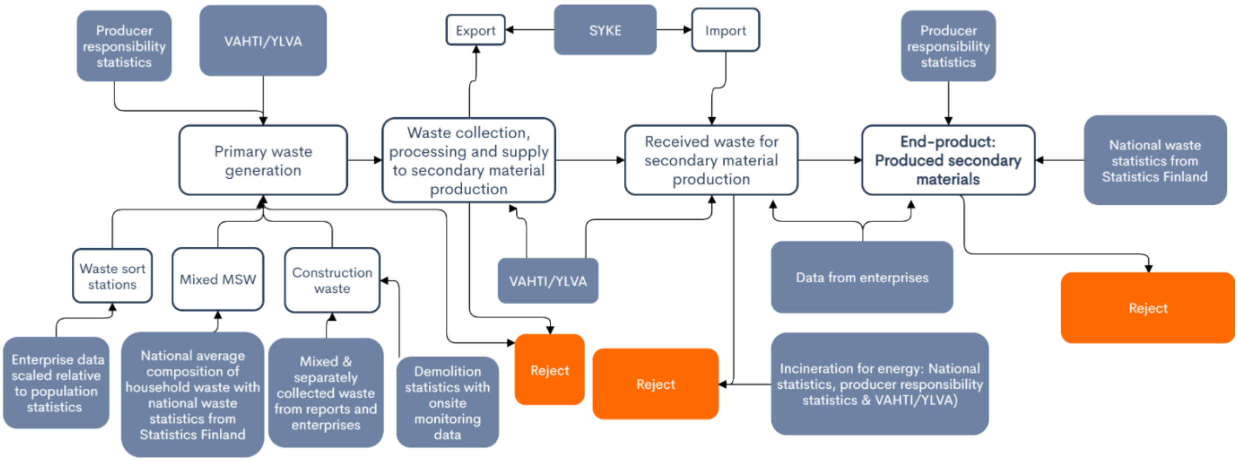

Figure 8: Data sources regarding wood in blue notes and orange notes indicating biggest data gaps/missing data.

The amount of *wood waste generated* could not be estimated with confidence due to existing severe gaps in data availability (Figure 8). Uncertainties regarding wood waste generated during demolition of buildings related to the inaccuracies in the demolition statistics as well as the missing data to assess the possible material combinations of the demolished buildings. Generation of more comprehensive data sets on the actual composition of demolished buildings would improve the accuracy of these estimates.

According to Häkämies et al. (2019) the amount of wood waste from construction and demolition was estimated to be approximately 250 000 t per year. Domestic supply in Table 13 combines this with the amount of wood packaging waste recycled as material and used for energy production according to producer responsibility statistics as well as the amount of wood waste brought to sorting stations. The latter was calculated by one company's collection rate and scaled to estimate the whole country by the ratios of company's collection area and Finland's total population.

The amount of wood waste in mixed MSW has been estimated to be 1,5% according to the national average composition of household waste (KIVO). In addition to this, there is also bulky wood waste coming from households, the amount of which could not be estimated here due to lack of data. The amount of wood waste in mixed C&DW was estimated to be 17-32%, for which 25% was used in this study (Liikanen et al. 2018). The average annual amount of waste generated from construction was estimated to be approximately 1,5 million tons of which 21% was estimated to be mixed waste and 15% wood (Salmenperä et al. 2016).

The data on *imported and exported* amount of wood waste were obtained from the Finnish Environment Institute SYKE's trans-boundary waste shipment statistics (SYKE).

Table 13: Merged supply table for wood waste in 2015 and 2018 in Finland.

Material	Domestic supply, t*	Import, t
Wood (2015)	528 373	48 000
Wood (2018)	541 092	30 000

* Collected wood waste

Wood waste rarely ended up in secondary material production and most were utilized in energy production, e.g. most of wood waste from demolition of buildings ends up in energy production (Häkämies et al. 2019). Produced secondary material amounts in Table 14 were obtained from producer responsibility statistics and include the repair of wooden pallets. Company inquiries failed to obtain additional information on the utilization of wood waste for secondary material production, the contacted companies utilizing industrial by-products rather than wood waste

Table 14. The use and production of secondary materials from wood waste in 2015 and 2018 in Finland.

Material	Export, t	Produced secondary material, t*
Wood (2015)	25 000	27 855
Wood (2018)	40 000	52 776

* Waste received for industrial manufacturing of secondary materials

2.3.8 Glass

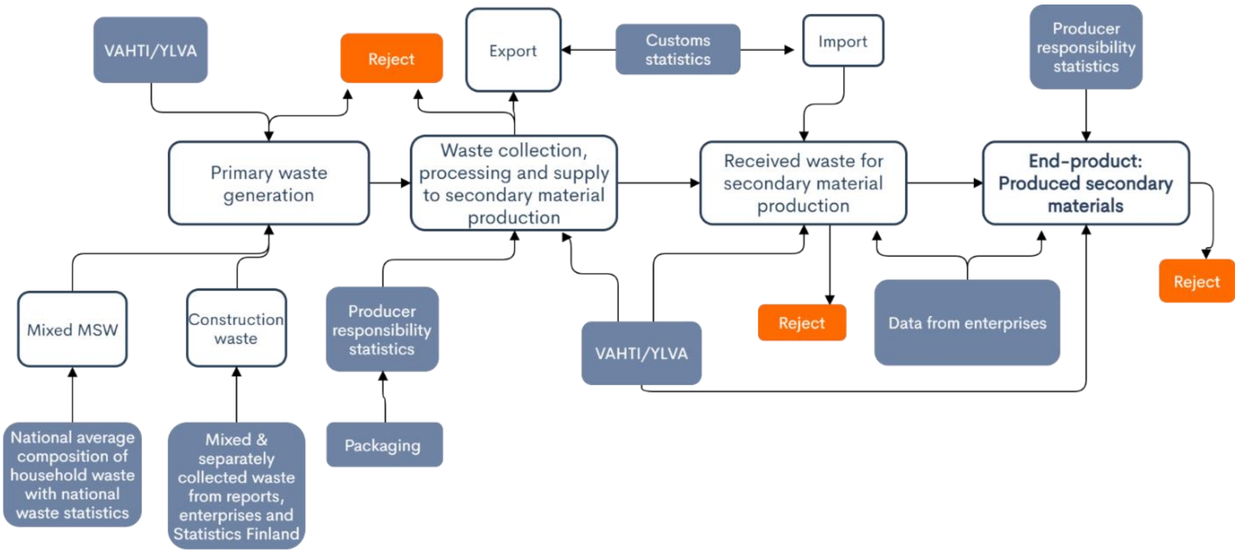

Figure 9: Data sources regarding glass in blue notes and orange notes indicating biggest data gaps/missing data.

Data on glass waste is based on data reported in VAHTI/YLVA, besides the import/export volumes of glass waste, which are based on customs statistics (Figure 9). Parallel data to VAHTI/YLVA was collected concerning glass waste from construction and demolition, and separately collected glass waste in MSW. The average annual amount of waste generated from construction was estimated to be approximately 1,5 million tons of which 21% is mixed waste and 1% is glass (Salmenperä et al. 2016). 3-6% of mixed construction and demolition waste was estimated to be glass waste, from which 5% was used in this study (Liikanen et al. 2018). The data regarding separately collected glass waste were obtained from Statistics Finland. Glass sorting has been moderately efficient and very little ends up in mixed household waste. According to the national average composition of mixed MSW from households, approx. 2% of mixed MSW consists of glass (KIVO), which was concluded to end up in energy production. Based on these parallel estimates, the domestic supply of glass waste (representing collected waste from MSW, construction and demolition) was estimated as 115 565 tons in 2015 and 120 735 tons in 2018, respectively (Table 15).

Table 15. The summary supply table for glass waste in 2015 and 2018 in Finland.

Material	Waste generation[1]	Waste collection and treatment[2]	Import, t[3]
Glass (2015)	12 103	69 371	22
Glass (2018)	25 463	66 229	22

[1]Generated glass waste reported in VAHTI/YLVA. Does not include waste pretreatment or storage.
[2]Waste reported as received by Waste collection, Waste treatment and disposal, and Materials recovery industries in VAHTI/YLVA. Does not include waste pretreatment or storage.
[3]Customs statistics

The amount of *secondary materials produced* (Table 16) were obtained from VAHTI/YLVA together with data regarding the recycled material from packaging waste (both domestically and abroad) from producer responsibility statistics.

There were no data available to estimate the amount of *reject* during waste collection, processing, supply to secondary material production and in receiving and producing secondary materials.

Table 16. The use and production of secondary materials from glass waste for years 2015 and 2018 in Finland.

Material	Export, t	Energy production, t	Waste to secondary material production, t*	Produced secondary material, t
Glass (2015)	47 109	35 171	88 014	55 379
Glass (2018)	32 324	30 438	90 312	73 114

*Waste reported in VAHTI/YLVA as received by the companies producing glass-based secondary materials operating in Finland

Glass packaging and other glass waste provided the only categories of glass in VAHTI/YLVA database suitable for analysis as waste consisting of only glass.

More efficient sorting of both MSW and C&DW waste could further improve the efficiency of recycling and increase the amount of produced secondary material from glass waste.

Concerning glass, we provide a set of more accurate supply and use tables for the years 2015 and 2018, in which the contribution of different industries to each phase in the glass waste-to-secondary material -cycle is demonstrated. These tables are provided in Annex I of the present document.

2.3.9 Mineral fractions from processed slag from waste incineration

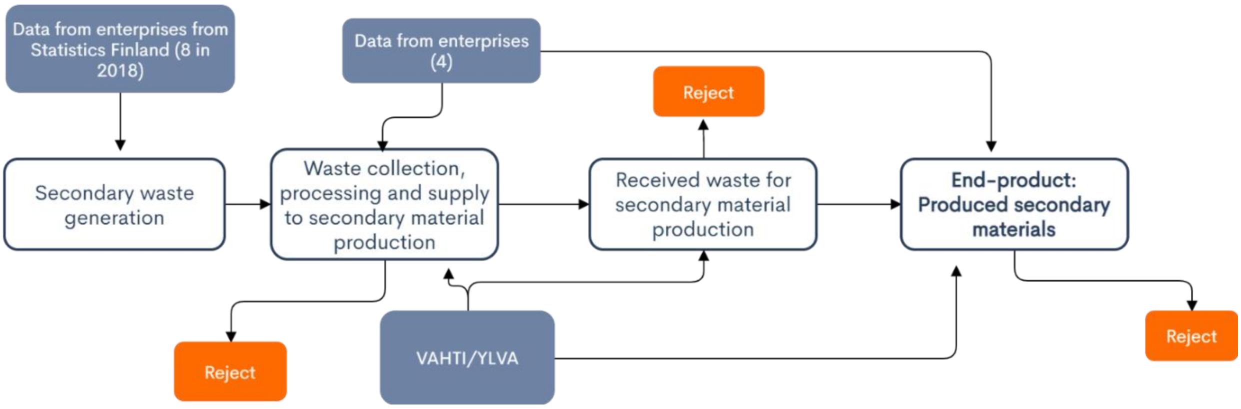

Figure 10: Data sources regarding mineral fractions from processed slag from waste incineration in blue notes and orange notes indicating biggest data gaps/missing data. In parenthesis the number of companies from which information was obtained.

The data on the *generation* of this secondary waste (i.e. waste generated from waste treatment) mineral fractions from processed slag from waste incineration, were obtained from incineration plants for year 2018 (Figure 10). The 2015 data were incomplete in the absence of data from companies. Uncertainty about the data was related to e.g. incinerated material and whether the incinerated material was defined as waste.

The data regarding the amount of slag treated by companies were obtained for both years in question (Table 17). The metals were either recovered from slag or before incineration. Therefore, the data collected from the companies could not be considered complete and comprehensive waste statistics requires information from both waste generation and treatment companies.

Table 17: Merged supply table for mineral fractions from processed slag waste in 2015 and 2018 in Finland.

Material	Domestic supply, t
Mineral fractions from processed slag, waste incineration (2015)	198 417
Mineral fractions from processed slag, waste incineration (2018)	285 981

According to our analysis, slag was not being *imported* and all supplied material flow was of domestic origin. The volume of waste incineration slag and ash exports in 2015 were obtained from customs statistics. According to the statistics there were no exports in 2018.

In the VAHTI/YLVA database, all slag and ash from any kind of thermal processing were reported in the same aggregated class, causing error in the estimates and making it impossible to distinguish the amount of slag from waste incineration that was the subject of this project.

Table 18: The use and production of secondary materials from mineral fractions from processed slag from waste incineration in 2015 and 2018 in Finland.

Material	Export, t	Produced secondary material, t
Mineral fractions from processed slag, waste incineration (2015)	1 271	117 201
Mineral fractions from processed slag, waste incineration (2018)	0	278 295

For the summary use and production table (Table 18), only the amount of produced secondary materials were available. Based on our analysis, the amount of produced slag-based mineral fractions was circa 117 200 tons in 2015 and 279 000 tons in 2020, with a larger data gap in the 2015 data. The difference between the domestic supply and produced secondary material were caused by storage. Other categories are omitted due to lack of data.

2.3.10 Fertilizers

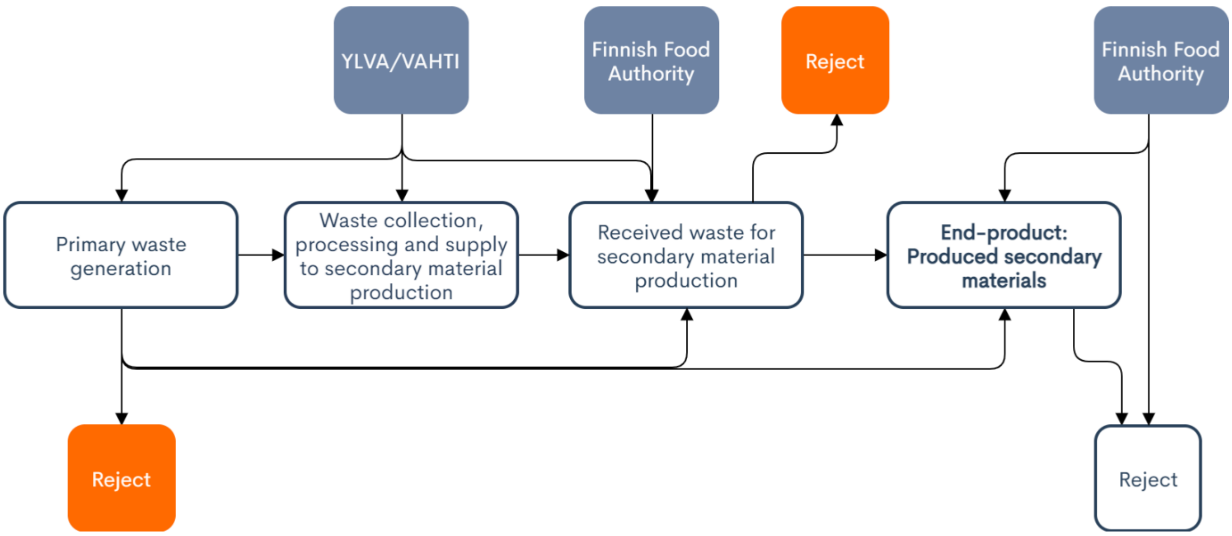

Figure 11: Data sources regarding fertilizers in blue notes and orange notes indicating biggest data gaps/missing data.

Recycled fertilizers are divided into two main groups: ash fertilizers and organic fertilizers (including solid and liquid ones). Finnish Food Authority regulates the production of fertilizers and is the main data source for recycled fertilizers. Figure 11 combines data sources for both ash and organic fertilizers.

Ash fertilizers consist of ash from wood and peat combustion. Data on *primary waste generation* includes all combustion wastes, but only some fraction of waste is used in fertilizer processing; most of it is unsuitable for fertilizer use and is backfilled or landfilled. YLVA/VAHTI includes data for waste generation and following steps of secondary material production. Additionally, Finnish Food Authority registry contains data on received ash waste for secondary material production. Most of the useful combustion waste is regarded directly as secondary material end-product, but ca. 20-25% of end-products are produced by manufactures of recycled fertilizers. Registry data of Finnish Food Authority includes detailed information of end-product use categories (agriculture, forestry, households, landscaping and maintenance of green spaces, other uses, export). Also end-products not used as fertilizers are reported.

Organic fertilizers are produced by several types of companies, only minority of them utilizing waste flows. Organic fertilizers are commonly by-products of some other production (goods or services). Sludges, animal and mixed food wastes, and manure are used in the production of solid organic fertilizers. Mixed and undifferentiated materials, animal feces, urine and manure, and vegetal wastes are most important waste flows in the production of liquid organic fertilizers. Biogas and potato starch plants are most important producers of liquid organic fertilizers. Only small fraction of suitable primary wastes is used in the production of organic fertilizers. The estimation of the amount of useful wastes with confidence was not possible due YLVA/VAHTI data gaps.

Table 19: Merged supply table for ashes from combustion of wood, peat, coal etc. in 2015 and 2018 in Finland.

Material	Domestic supply, t
Combustion ashes (2015)	1 394 868
Combustion ashes (2018)	1 409 327

The supply of combustion waste was exclusively domestic in 2015 and 2018 (Table 19). The supply includes all ash and slag, most of it unsuitable for recycled fertilizers. Domestic supply for organic fertilizers was not possible to estimate due to the large heterogeneity of waste flows and manufacturing processes of solid and liquid organic fertilizers.

Table 20: Use of combustion ashes as raw material for secondary fertilizer production, supply of ash, solid organic and liquid organic fertilizers and related volumes of reject in 2015 and 2018 in Finland.

Material	Raw materials for secondary material manufacturing, t	End-product, Produced secondary material, t[1]	Rejected end-product, t[2]
Ash fertilizers (2015)	45 324	29 165 (121 656*)	45 833
Ash fertilizers (2018)	86 008	55 813 (176 548*)	96 240
Solid organic fertilizers (2015)	-	14 641	0
Solid organic fertilizers (2018)	-	19 944	0
Liquid organic fertilizers (2015)	-	152 657	9 728
Liquid organic fertilizers (2018)	-	75 932	15 899

[1] First figures of ash fertilizers are end-products from secondary material manufacturing. Figures with asterisk (*) represent the amount of secondary materials produced directly without processing phase. Sum of two figures row wise = total amount of end-products.
[2] End-product not used as fertilizer (backfilling, losses etc.) from total produced secondary material.

Production of secondary materials differs between ash and organic fertilizers (Table 20). Some amount of useful combustion waste act as raw materials for secondary material manufacturing. Most of the waste are considered as end-product without any treatment or manufacturing. 30-40 % of combustion wastes qualified as ash fertilizers are not used as fertilizers (regarded as rejected in Table 20). Organic fertilizers don’t have a similar production path, hence there is no raw materials data available. Some of organic fertilizers are also rejected.

2.3.11 Sludge from wastewater treatment

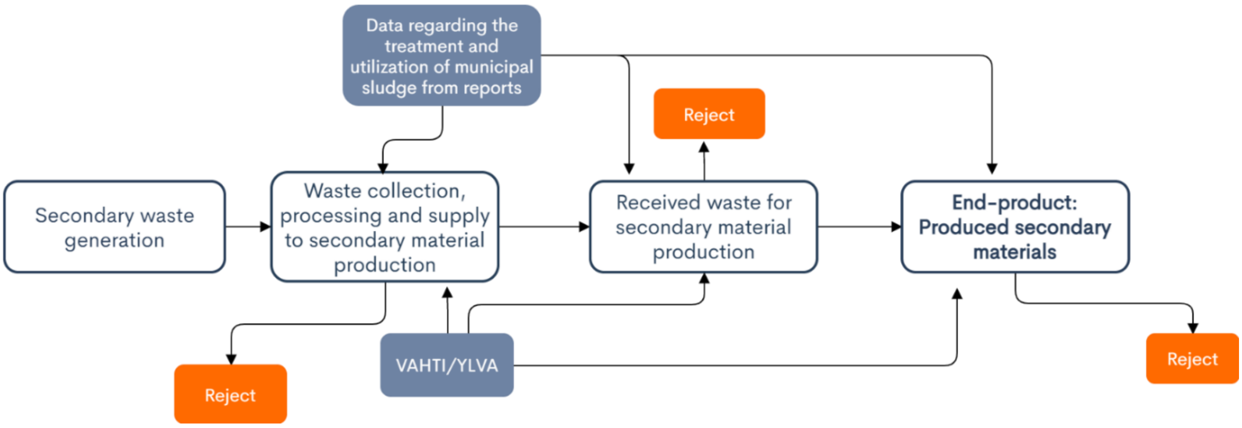

Figure 12: Data sources regarding sludge from wastewater treatment in blue notes and orange notes indicating biggest data gaps/missing data.

Data regarding the treatment and utilization of the sludge from municipal wastewater treatment in 2015 and 2018 (Figure 12) were obtained from two comprehensive studies done in 2017 and 2019 by Finnish Water Utilities Association (FIWA) (Vilpanen & Toivikko 2017; Konola & Toivikko 2019). The studies supplemented the VAHTI/YLVA data with, for example the Finnish Food Authority's annual declarations data, the monitoring data of the ELY Centers and direct enquiries to operators. According to the studies sludge is mostly utilized and only rarely incinerated in Finland. Data were not available from all the treatment plants, one reason being that data regarding treatment plants that have a municipal environmental permit were not readily available. Better availability of these permits would improve the accuracy of the data.

The data regarding wastewater sludges from industrial processes were obtained from VAHTI/YLVA database. Some estimates for the ratios of sewage to process sludge could be obtained through business surveys or closer studies on the subject. To exemplify, in 2018 wood, paper and pulp, and food industries generated roughly 140 000 tons of organic (rather than inorganic) sludges but the composition and fate of these wastes is very heterogenous and therefore needs additional investigation in future (Table 21).

Table 21: Total supply and supply to material recovery of sludge from municipal wastewater treatment plants in 2015 and 2018 in Finland.

Material	Total supply, t	Material recovery, t
Sludge from wastewater treatment (2018)	133 567	112 687
Sludge from wastewater treatment (2018)	173 917	144 325

2.3.12 Sand, gravel, clay and crushed stone

For the recovery of sand, gravel, clay and crushed stone, data was not available. This is largely due to the fact that the government decree on the recovery of soils classified as waste did not come into force as expected and hence data on this material and its recovery is not available at all. This decree is expected to improve the data availability significantly. The available data is limited to crushed stone but as this data is not comprehensive, it is not presented.

VAHTI/YLVA contains data on contaminated soils that have been received at waste management facilities having a permit for their treatment. Such soils are, however, often landfilled and no data are available on the fraction that is treated and used for earth works elsewhere. In addition, in this database all these types of waste are aggregated into one single waste class, meaning that it is not possible to assess the contribution of different types of soil. Finally, assessing amount of reject, export, energy production or end-product via this data source alone is not possible. According to the data received from Confederation of Finnish Construction Industries RT, the volumes of aggregates reused within a construction site or received from other sites were up to 67 Mt and 74 Mt in 2015 and 2018, respectively. These material flows, however, cannot be classified as secondary materials in similar manner as for the rest of the materials in the report. For the above reasons, summary accounts for secondary sand, gravel and clay were not compiled.

2.4 Discussion and conclusions

In the current study, we estimated the material flows relating to secondary material production in Finland and developed methodology for compiling such accounts. These secondary material accounts can be used in compiling material-specific supply and use tables, an example of which was provided for glass waste and its use for the production of secondary glass materials.

Data availability, quality and coverage varied greatly from material to material and according to the stage of the waste-to-secondary material -cycle. Overall, the most reliable sets of data were collected for the last step of the cycle, that is, the actual manufacture of the secondary materials. For paper, cardboard, different metals, plastics, glass, mineral materials from incineration of slag, and fertilizers, we assess that the supply of the secondary materials were highly representative for the Finnish economy (particularly in 2018). For crushed concrete and bricks, these pieces of the accounts were somewhat less reliable due to the lower coverage of the primary data.

For accounting purposes, the data on the earlier stages of the waste-to-secondary material -cycle contained considerable shortcomings. In the accounts, the supply and use should always be equal with minimal balancing items. E.g. supply of a specific waste from other industries should be equivalent to the use of this particular waste by the secondary material manufacturing industries. In this work, these figures were compiled independently using different and varying sources of data. This way, we estimated the data availability and quality for each stage of the cycle independently. To balance the accounts e.g. supply and use at different stages would need additional efforts. To this end, the representativeness of the data at each stage would need to be assessed. In many cases, however, this would be laborious if not impossible considering the current availability of data e.g. on the industry-specific generation of waste.

The case of glass, which was used as an example of the compilation of the actual accounts in subchapter 2.3.8, sheds light on the problem at hand. Data on waste generation was obtained from VAHTI/YLVA compliance monitoring system but they represent only facilities with an environmental permit granted by a state authority. In principle, these data could be used as a basis for industry-specific upscaling based on revenues, economic outputs or manufacturing volumes similarly to the accounting work by Salminen et al. 2018 and Weckström et al. 2020 for water and wastewater accounts. However, unlike water use and wastewater generation, glass waste generation in manufacturing facilities may often arise from other than regular core operations of a facility. This was suggested by the very different industry-specific glass waste generation volumes in 2015 and 2018 (see Annex I). Hence, such generalization seems methodologically invalid and would likely result in biased results. Also, waste glass collected via the deposit refund system (Suomen palautuspakkaus Oy Palpa) generates one aggregate figure containing beverage bottles returned from households and all other sectors of the economy. Hence, no industry-specific data for this figure can be calculated. This applies to glass waste collection arranged by waste management industry in general.

In the course of the work several potential avenues to facilitate and to improve the compilation of high-quality secondary material accounts were observed. First, data reporting in VAHTI/YLVA compliance monitoring system should be made unambiguous in terms of waste nomenclature, generation, reception, treatment and recovery and storage of waste and choices related to that. Risks of double counting are considerable and for their elimination plenty of manual data reviewing is required. In addition, great variation exists in the way enterprises report their data on the generation and use of waste to VAHTI/YLVA system. This hampers the extractability ja usability of the VAHTI/YLVA data for accounting and research purposes. Introduction of identifiers for individual sets of waste would be useful to allow their tracking along the multiple stages of waste-to-secondary material -cycle. The revision of the waste information system is underway in Finland. Some improvements e.g. to the higher accuracy in the origin of waste are expected. The considerations presented here related to secondary material accounting merit to be addressed in the revision process.

New, voluntary systems (databases) benefitting from sustainability reporting by individual companies and organizations or linked to green deals made by industries might also improve the availability and accessibility of data that is essential to the compilation of the accounts along the waste-to-secondary material -cycle. For now, according to the experience gained in ENVECO project, compilation of secondary material accounts is laborious, time-consuming, and prone to biases.

Economy-wide secondary material accounting benefits greatly from disaggregated data both by origin of waste and its material composition, since it enables more efficient targeting and monitoring of policies related to waste and CE. The current waste statistics and material flow accounts tend to categorize waste types and provide data in an aggregated form, making it more difficult and uncertain to utilize the data in accurate, material-specific waste accounts. Taking the need for disaggregation into account in development of reporting schemes would make it more reliable to utilize the data in high-resolution accounts. Utilizing smart digitalization could enable more detailed accounting without excessive administrative burden.

In addition, there is great variation in the way enterprises report the generation and use of waste. Data management within each company affects how easily the data is available to be extracted for accounting and research purposes. Often data on waste is still scattered and difficult to access, even within companies.

2.5 References

Alenius, M. 2016: Advancing the recycling of agricultural plastics. Case: Envor Group Oy.

Bachelor's Thesis in Environmental Engineering. 46 pages, 3 pages of appendices. Lahti University of Applied Sciences, Degree Programme in Environmental Technology.

Eurostat (2018). Economy-wide material flow accounts handbook. 2018 edition. European Union.

Erälinna L, Järvenpää AM. 2019: Agricultural Plastic Waste Collection and Recycling - Challenges and Opportunities. University of Turku, Finland.

Forstén L (2015). Asfaltti – monikäyttöinen ja uudistuva [Asphalt – a versatile and modernizable material]. A seminar presentatioFors at Infra-alan ympäristöpäivä [Environment day of infra industries], 23.10.2015, Helsinki (In Finnish, seminar and presentation title translations in brackets by authors of this report)

Häkämies, S, Lehtonen K (2019). Puupohjaisen rakennus- ja purkujätteen määrä ja laatu, jakautuminen eri alajätejakeisiin sekä syntypaikat [The quantity, quality, constituents and sites of generation of wood-based C&DW]. A seminar presentation at Puu rakentamisen kiertotaloudessa [Role of wood in CE concerning the construction industry] event, 19.9.2019, Helsinki (In Finnish, seminar and presentation title translations in brackets by authors of this report).

Konola I., Toivikko S. (2019). Yhdyskuntalietteen käsittelyn ja hyödyntämisen nykytilannekatsaus. [Current status on processing and utilization of municipal sewage sludge]. Vesilaitosyhdistyksen monistesarja nro 57. Finnish Water Utilities Association (FIWA). (In Finnish, seminar and presentation title translations in brackets by authors of this report).

Liikanen, M., Helppi, O., Havukainen, J., & Horttanainen, M. 2018. Rakennusjätteen koostumustutkimus – Etelä-Karjala. LUT Scientific and Expertise Publications –Research Reports 82. Lappeenranta University of Technology.

Salmenperä, H, Sahimaa O, Kautto P, Vahvelainen S, Wahlström M, Bachér J, Dahlbo H, Espo J, Haavisto T, Laine-Ylijoki J (2016): Policy instruments for increasing waste recycling. Publications of the Govenrment´s analysis, assessment and research activities 53/2016

Seppälä et al. (2011) An assessment of greenhouse gas emissions and material flows caused by the Finnish economy using the ENVIMAT model. J. Cleaner Prod. 19: 1833–1841

Nylén EJ, Salminen J. (2019) How does the circular economy discourse affect policy-making? The case of streamlining waste utilisation in Finnish earthworks. Resources, Conservation & Recycling 149: 532-540.

Salminen JM, Veiste PJ, Koskiaho JT, Tikkanen S. (2018) Improving data quality, applicability and transparency of national water accounts – A case study for Finland. Water Resources and Economics 24: 25-39.

Vilpanen, M, Toivikko, S. (2017). Yhdyskuntalietteen käsittelyn ja hyödyntämisen nykytilannekatsaus. [Current status on processing and utilization of municipal sewage sludge]. Vesilaitosyhdistyksen monistesarja nro 46. Finnish Water Utilities Association (FIWA). (In Finnish, title translation in brackets by authors of this report).

Weckström M, Örmä V, Salminen J. (2020) An order of magnitude: How a detailed, real-data-based return flow analysis identified large discrepancies in modeled water consumption volumes for Finland. Ecological Indicators 119: 105835

3 Application of machine learning, remote sensing and spatial analysis methods in forest ecosystem indicators (WP3)

Pekka Hurskainen[2], András Balázs[1], Janne Mäyrä[2], Sakari Tuominen[1], Annika Kangas[1], Petteri Vihervaara[2]

[1] National Resources Institute Finland (LUKE)

[2] Finnish Environment Institute (SYKE)

Chapter 3 addresses the possibilities of developing indicators measuring forest ecosystem condition within SEEA-EEA framework in Finland, with support of machine learning, remote sensing, and spatial analysis methods.

The chapter starts with a brief outlook and discussion on ecosystem accounting indicators in SEEA-EEA framework (chapter 3.1) and literature review on forest ecosystem condition indicators, including assessment of possibilities for producing wall-to-wall maps of these indicators with remote sensing and existing spatial data. This was the **first objective** of WP3.

Chapter 3.2 serves as an introduction to current state-of-the-art on machine learning methods in remote sensing of forests.

The **second objective** of WP3 was initially to test if introducing machine learning methods, such as deep learning, will improve the prediction results of selected forest variables and indicators compared to traditional RS modeling methods presently used in Finnish multi-source national forest inventory (MS-NFI). However, as we had completed the literature review on indicators, we saw it equally appropriate and feasible to look at existing national spatial datasets (MS-NFI and topographic map database) and see how they could be used as source data for ecosystem condition indicators. Specifically, we wanted to find out how forest fragmentation could be quantified and monitored with landscape indicators and application of existing geospatial analysis tools and methods. For both approaches, data and methods are elaborated in Chapter 3.3 and results and discussion in Chapter 3.4.

Assessment of uncertainties related to these methods, our **third objective**, is also discussed in Chapter 3.4. Conclusions and recommendations for further research are given in Chapter 3.5.

The **fourth and last objective** of WP3 was to provide the value of forests by using monetary asset accounts. The value for timber production was calculated by using three approaches. The value of carbon sequestration was expressed as net present value for different prices of carbon. Since the forest valuation part is different in scope and methodology, it is discussed separately in Chapter 4.

3.1 Indicators in ecosystem accounting

The role of the SEEA EEA is to integrate data from various sources such as basic statistics, surveys, scientific measurements, spatial and remote sensing data, administrative entities, and censuses, to provide a coherent and unified understanding of ecosystems and their relationship to the economy. Information from the various ecosystem accounts and related accounts in the SEEA EEA framework can be organized and integrated to provide policy-relevant indicators and aggregates (UNSD 2020). According to the draft SEEA-EA guidelines, indicators are defined as follows:

> "A statistical indicator is the representation of statistical data for a specified time, place or any other relevant characteristic, corrected for at least one dimension (usually size) so as to allow for meaningful comparisons. It is a summary measure related to a key issue or phenomenon and derived from a series of observed facts. Indicators can be used to reveal relative positions or show positive or negative change in a regular interval. Indicators are usually a direct input into national and global policies. In strategic policy fields, they are important for setting targets and monitoring their achievement. By themselves, indicators do not necessarily contain all aspects of development or change, but they greatly contribute to explaining them. If consistent methodology is employed, they allow comparisons over time and between, for instance, countries and regions, and in this way assist in gathering 'evidence' for decision making".

Furthermore, three types of indicators are considered within the draft SEEA-EA (UNSD 2020), namely

a) **Aggregates**, statistics for related categories that can be grouped together or aggregated in order to provide a broader picture. Example: summing the areas of different forest ecosystem types across an ecosystem accounting area.

b) **Composite indices,** in which different variables are combined using a weighting pattern or aggregation rule to communicate the overall movement or trend. Example: measures of forest ecosystem condition which involve weighting together relevant condition indicators.

c) **Ratio indicators,** derived by combining data from different accounts. Example: the flows of ecosystem services per hectare from different forest types.

These indicators can be derived from any of the core accounts (ecosystem extent, ecosystem condition, ecosystem service supply and use in physical and monetary terms, ecosystem asset account) or thematic accounts (biodiversity, climate change, oceans, urban areas) of the accounting framework. Sometimes the same indicator can be even used in several accounts with slight modifications.

In Finland, the implementation of SEEA-EEA framework, as well as reaching a consensus of commonly agreed indicators, are still a work in progress. The development of National Ecosystem Services Indicators (www.biodiversity.fi/ecosystemservices/home; see Mononen et al. 2016 and Lai et al. 2018) following the CICES classification system and the Cascade model (Haines-Young and Potschin 2010) has had a good start. However, further work is needed especially for the ecosystem condition indicators, in leveraging the possibilities of remote sensing, machine learning and spatial analysis methods, and practical implementation to bring these indicators to the accounting framework.

Assessment of ecosystems and their services are set in the EU Biodiversity strategy to 2020, involving assessment of ecosystem condition, since condition directly impacts ecosystem capacity to provide ecosystem services (Maes et al. 2013, 2014). In order to be able to account for ecosystem services, we understand that the extent and condition of ecosystem assets should be known and accounted for first, while taking into consideration what is feasible and realistic with the source data, methods and resources available.

In this study, we are concentrating on ecosystem condition indicators for forest ecosystems.

3.1.1 Indicators on forest ecosystem condition

The SEEA ecosystem condition typology (ECT) is a hierarchical typology for organizing data on ecosystem condition characteristics (UNSD 2020). The ECT contains three groups and six classes. The first two classes (physical and chemical state) describe *abiotic ecosystem characteristics*, forming the first group. The next three classes (compositional, structural and functional state) describe *biotic ecosystem* *characteristics*, the second group, and the sixth class describes *landscape level characteristics* which is the third group. The condition indicators are meant to be calculated for each ecosystem type and disaggregated by ecosystem condition subclasses (UNSD 2020).

In this project, we scanned the existing literature on published ecosystem condition indicator sets, and selected the set published in Czuz et al. (2018a) and Maes et al. (2018), as comprehensive and relevant for our purposes. We then reviewed the indicators and re-organized the table to follow the ECT in order to find a subset that is both comprehensive (i.e. suitable for describing forest ecosystem condition in Finnish boreal forests) and either observable with remote sensing or derivable from existing spatial data (Table 22 and Annex 2). The most promising indicators from this subset were finally selected for further testing and development. The indicators tested and developed here can be classified as structural state and landscape indicators following the ECT.

Table 22: Indicators on forest ecosystem condition and the potential of Remote Sensing and existing spatial data (e.g. national forest inventories, topographic map database, etc.)

SEEA ECT class	Description and examples of indicators	Potential of Remote Sensing	Potential of existing spatial data	Indicators tested/developed in this study
Physical state indicators	Physical characteristics of forest asset abiotic components, such as soil structure (e.g. soil organic carbon, soil thickness) and water availability	Low	Low	-
Chemical state indicators	Chemical composition of the forest asset abiotic components, such as nutrient / material balances (e.g. acidification, chemical status of ground water, tropospheric ozone concentration)	Low	Low	-
Compositional state indicators	Characteristics of forest asset compositional state, such as composition / diversity of ecological communities at a given location and time (e.g. invasive alien species richness, forest tree species richness, bird indices, deadwood, naturalness)	Medium	Medium	Forest tree species groups distribution
Structural state indicators	Overall structural state characteristics of a forest asset, including characteristics on vegetation, biomass, and food chains (e.g. NDVI, Leaf Area Index, Growing stock biomass volume, structural heterogeneity)	High	High	Forest stand indicators: total volume, volume by tree species groups, diameter at breast height, average tree height
Functional state indicators	Overall functional state characteristics of a forest asset, including characteristics on ecosystem process and disturbances regimes (e.g. primary productivity, forest age, community age, disturbance frequency)	Medium	Medium	Average tree height as a proxy for age
Landscape indicators	Overall characteristics of forest ecosystem assets that are quantifiable at larger (landscape) scale but that have an influence on the local condition of ecosystems (e.g. fragmentation, connectivity).	Medium	High	Fragmentation patterns of natural/seminatural landscapes, fragmentation patterns of forest landscapes

It should be also noted that ECT include only indicators describing **ecosystem state.** This reflects the ‘state orientation’ of condition indicators as a preference (Czúcz et al. 2018c). Alternatively, when little data on state is available, ecosystem condition can also be measured through **ecosystem pressures** as an indirect approach (UNSD 2020). This should be done with caution, however. Considering that pressures lead to degradation (depletion, accumulation) of environmental stocks, it is the stocks themselves that should be considered as ecosystem condition variables instead of their change (degradation / depletion rates, fluxes, flows) (Czúcz et al. 2018c). State indicators should be preferred as they can be used to formulate clear policy messages on ecosystem degradation; and the degree of policy attention highlights those environmental stocks that are perceived as the most valuable or most endangered (UNSD 2020). For these reasons, in this study we consider only state indicators, not pressures.

A more extensive annotated table including an exhaustive list of proposed forest ecosystem condition indicators can be found in Annex 2. Like stated earlier in this chapter, this table is a mostly based on existing and exhaustive work done in Czúcz et al. (2018a) and the 5th MAES report (Maes et al. 2018). Our additions to this table in Annex 2 include re-organizing it to follow ECT and two additional columns: potential of remote sensing and potential of MS-NFI, which give our descriptive assessment of the potential in producing these indicators in Finland with remote sensing methods and spatial analysis of multi-source national forest inventory data. After all, if the objective of ecosystem accounting is to be spatially explicit with the ability to produce both tables and maps from these accounts, we need to have also spatially explicit source data for the indicators.

3.2 Literature review: Machine learning methods in the remote sensing of forests

3.2.1 Background

Remote sensing of forests is typically multivariate analysis of digital remote sensing data, containing values of satellite and aerial image pixels or lidar and photogrammetric point clouds. In addition to the values of individual pixels or points, the structure and organization of neighboring pixels and points are a basis for features that are used to predict the characteristics of forests. The features that can be extracted from remote sensing data include (but are not limited to) spectral tone (color), texture, pattern, shape, shadows or association (to other features) (Lillesand et al. 2004). E.g. spectral tone can be applied to classifiers or estimators based on e.g. regression or similarity in spectral feature space. The more complex features are applied, the higher learning capability is required of the classifiers and estimators. When utilizing high-resolution multi- or hyperspectral image data in combination with geometric point clouds, sufficient machine learning capability is required in the data analysis in order to utilize the full potential of the data.

Nowadays, various machine learning algorithms are the method of choice for processing large data sets with complex, nonlinear structure. The most widely used ML methods in remote sensing are support vector machines (SVM) and random forests (RF) due to them working well with high-dimensional data, being straightforward to use and not requiring heavy computational resources. Both methods have been used for multiple tasks, such as predicting land cover (e.g. Gòmez et al (2016)) and classifying tree species from Sentinel-2 satellite data (Grabska et al (2020)) and hyperspectral remote sensing data (Viinikka et al (2020), Modzelewska et al (2020)). Aside from SVM and RF, other commonly used ML algorithms are k-nearest neighbors (kNN), gradient boosting trees (GBM), gaussian processes and various deep learning (DL) methods.

In the following we take a closer look at a number of potential machine learning methods for the remote sensing of forests based on literature review.

3.2.2 K nearest neighbors method

K-nearest neighbors algorithm (k-NN) is a machine learning method which is widely applied in remote sensing based forest inventory. It was originally developed by Fix and Hodges (1951). It can be used both for classification and estimation of continuous numerical variables. In both cases the neighbors are taken from a set of objects for which the class (for k-NN classification) or the property value (for k-NN regression) is known. This can be considered as the training set for the algorithm, although no explicit training phase is required.

The method was first suggested for satellite image-based forest inventory in the form of reference sample plot method by Kilkki & Päivinen (1987), and subsequently it has been widely applied in remote sensing based forest inventories (Muinonen and Tokola 1990; Tomppo 1991; Tokola et al. 1996, Tokola & Muinonen 1996, Nilsson 1996, Franco-Lopez et al. 2001). .At the moment, k-NN is the method used e.g. in Multi-Source National Forest Inventory (MS-NFI) to predict wall-to-wall maps of forest parameters.

In the k-NN method the training consists of the k nearest observations in feature space. The method has several advantages: it is non-parametric, not requiring parametric models for the estimation or classification task, all output variables of interest can be determined at the same time, while simultaneously retaining the covariance structure of e.g. forest variables of training data. Additionally, as the method does not extrapolate values outside the training data, it will not produce unrealistic estimations in the case when independent test data feature space is different from training data.

K-NN is called instance-based learning, or also lazy learning, because instead of performing explicit generalization, the 'model' is approximated only locally; the algorithm compares new observations with observations present in the training data. The algorithm generally works well in multi-dimensional feature spaces with relatively low number of dimensions (such as satellite image bands), although training data features representing different physical units or markedly different scales of variation need to be normalized for optimal performance.

The main drawback of the method is related to the fact that current state of the art remote sensing data sets, such as combinations of lidar or photogrammetric point clouds and high-resolution imagery data, make it possible to derive a very large number of remote sensing features describing the forest characteristics. Consequently, the dimensionality of the remote sensing feature space increases greatly, and it may be computationally infeasible to use all possible remote sensing features when processing large inventory areas. Also, when the dimensionality grows, all training data become sparse in relation to the dimensions (Hinneburg et al. 2000), which makes estimators that are based on distance or proximity in the feature space, such as k-NN, prone to 'the curse of dimensionality’ (Beyer et al. 1999); the distances in the feature space become insignificant, when the number of dimensions exceed the number of training observations.

Thus, higher-level machine learning algorithms are required for producing an appropriate subset of features or appropriate weights for a number of features for the estimation procedure, considering their usefulness in predicting the forest attributes as well as their mutual correlation. In the following we present number of alternatives for this purpose.

3.2.3 Sequential forward and backward selectors

Sequential forward and backward selection methods (Whitney 1971) are based on iterative selection of RS features for picking an optimal data set for the estimation task. Sequential forward selection starts from empty set of data features, and in each consecutive iteration the feature giving the best estimation result with already selected features is selected. The sequential backward selection which is seldom applied in the context of remote sensing starts with a set containing all available features and, in each iteration, the worst performing feature is excluded from the set. In both methods the iteration proceeds until there is no improvement in the estimation result (Järnstedt et al. 2012).

Typically, the measure for the RS based estimation result is proportion of correct classifications, or root mean square error or bias of numerical variables estimated. Typical weakness of the sequential forward/backward selection methods are their tendency to get stuck in local optima, e.g. in the case when the addition of any extra feature does not improve the estimation result individually, the algorithm is not able to proceed further. E.g. based on Kudo & Sklansky (2000), testing several feature selection algorithms has shown that sequential forward/backward selectors work best in relatively simple estimation tasks and in feature space whose dimensionality is relatively low (such as image bands of medium resolution satellite imagery). On the other hand, it has been noted that even in the feature space formed by Landsat TM satellite image bands, sequential forward selection algorithm may get stuck with the first (best performing) image band, when adding any individual band does not improve the estimation or classification result, even if the use of all image bands provides better result than any band individually (Holopainen ym. 2009).

A variation for sequential forward selection that is called nested forward selection has been suggested for dealing with the method's liability for the local optima (Pohjankukka et al. 2018). in this variation the original data set is divided to N subsets, and the selection of data features is carried out from each subset. The objective of the method is to examine the sensitivity of the feature selection to changes in the input data set.

3.2.4 Random Forest

Random forest (RF) is a decision tree-based machine learning method for e.g. classification and regression tasks (Breiman 2001). It is based on constructing multiple decision trees that serve as models in predicting the variables on interest. RF has been found to require less complex computations and running time compared to most other classifiers, as well as having high classification accuracy in especially intricate models (Cutler et al. 2007). Typically, in the training phase the algorithm builds a multitude of decision trees, and the algorithm tests various decision trees based on subsets (resamples) of predictor variables for finding the optimal result for the variables of interest. In the decision trees each node represents a test on a variable, and each branch represents an outcome of the test.

RF algorithms have several operational advantages. It typically requires fewer parameters than other classifiers (Akar and Gungör, 2012). RF is multivariate, non-parametric and quick to implement; however, the method also has limitations. Since RF predictions are weighted averages of the training data, the algorithm cannot extrapolate beyond the range of the training data. RF methods also typically require a greater number of training observations than parametric regression to perform well. The main advantage of the RF in analyzing remote sensing data is their capability to model complex interactions between multivariate data.

An individual decision tree is always somewhat prone to overfitting within certain training data set, but this problem can be controlled by using sufficient number of decision trees (Stephens 2017). Generally, the use of RF method assumes sufficiently large training data set. RF method has been widely used method in remote sensing of forests (Ballanti et al. 2016, Tuominen et al. 2018, Sabat-Tomala et al. 2020).

3.2.5 Support Vector Machines

Support vector machines (SVM) are supervised learning models that are used for classification and regression tasks. A support-vector machine is used for constructing hyperplanes in multi- or hyperdimensional feature space, and these hyperplanes are used for the separation or classification of observations based their input values in the feature space. SVMs are used to define a space in which the different classes are maximally separable. The hyperplanes are adjusted in the feature space in the way that maximizes the margin separating the classes in the training data, and unlabeled samples are then classified according to the side of the hyperplane on which they fall. Linear SVMs are the most common form, also nonlinear variants have also been designed (Kumar et al. 2020). Good separation is achieved by a hyperplane that has the largest distance to the nearest training-data point of any class, since in general the larger the margin, the lower the generalization error of the classifier (Hastie et al, 2008). SVMs have been used for wide variety of data and applications. which can be used for classification, regression, or other tasks like outlier detection.

SVM is generally considered a well-suited machine learning method e.g. for image interpretation (Ballanti et al. 2016, Sabat-Tomala et al. 2020). The main advantages of SVM are its' applicability and good separation capability in high- or hyperdimensional feature space as well as the classification capability without a priori information. SVM is trained by training data, based on which the SVM can classify any independent test data. The representativeness of the training data determines the quality of the classification in independent data.

3.2.6 Genetic algorithms

The genetic algorithms a are based on imitating evolutionary natural selection. The machine learning procedure in remote sensing typically contains following phases: The general GA procedure begins by generating an initial population of strings (corresponding to chromosomes or genomes), that consist of random combinations of predictor variables (which correspond to genes). The chromosomes are considered binary strings having values of 0/1 indicating that certain variable is either "selected" in the subset or "not selected". The strings evolve over a user-defined number of iterations (i.e. generations). This evolution may include following operations: selecting strings for mating by applying a user-defined objective criterion (the more copies in the mating pool, the better), letting the strings in the mating pool to swap parts (cross over), causing random noise (i.e. mutations) in the offspring (i.e. children). The resulting strings are passed to the next generation. This process is repeated until a predefined criterion is fulfilled or a predetermined number of iterations have been completed (Broadhurst et al. 1997; Tuominen and Haapanen 2013; Moser et al. 2017).

Typical task for a genetic algorithm in remote sensing is the selection of an appropriate set of remote sensing features from a hyper-dimensional feature space. Additionally, other estimation and classification parameters such as e.g. value on k in k-NN estimation can be optimized by GA. Typical criteria for a GA procedure in remote sensing is for example minimizing the RMSE or bias of remote sensing based forest estimates, or maximizing the proportion of correct classifications.

As in natural processes, there is a random effect in a GA procedure, which affects the result. Additionally, user-defined parameters such as population size, number of generations, amount of randomness and mutations allowed affect the results of the procedure.

The main advantage of the GA method in remote sensing is its' proven performance in empirical estimation and classification tasks (e.g. Haapanen & Tuominen 2008, Tuominen et al. 2018). Although there is no guarantee of finding the subset representing a global optimum within certain input variable dataset, and the algorithm does not go through all possible combinations input variables (which often would be also computationally unfeasible), solutions close to optimal can usually be found in a feasible computational time by GA (Garey and Johnson 1979).

3.2.7 Artificial neural networks

Artificial neural networks (ANNs), often called simply neural networks are computing systems that try to imitate biological neural networks. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons.

An artificial neuron that receives a signal then processes it and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called edges. Neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold. Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traversing the layers multiple times.

Neural networks trained by processing examples, each of which contains a known "input" and "result," forming probability-weighted associations between the two, which are stored within the data structure of the net itself. The training of a neural network from a given example is usually conducted by determining the difference between the processed output of the network (often a prediction) and a target output. This is the error. The network then adjusts its weighted associations according to a learning rule and using this error value. Successive adjustments will cause the neural network to produce output which is increasingly similar to the target output. After a sufficient number of adjustments, the training can be terminated based upon certain criteria. This is known as supervised learning.

Such systems "learn" to perform tasks by considering examples, generally without being programmed with task-specific rules. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any prior knowledge of cats, for example, that they have fur, tails, whiskers, and cat-like faces. Instead, they automatically generate identifying characteristics from the examples that they process.

Multilayer perceptron (MLP) (Bishop 1995) is a simple form of ANN. It is feedforward NN, I.e. the information moves in only one direction, that is forward from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network where the connections between the nodes do not form a cycle, as in recurrent neural networks. The feedforward neural network was the first and simplest type of artificial neural network devised. Typically, MLP NN is composed of input layer, one or more hidden layers and output layer. The purpose of MLP is to train the NN to minimize the difference between NN output values and the correct values picked from the training data, by adjusting the weighting between the inputs. Thus, the training of MLP is multivariate optimization task, aiming at minimizing the error of user-defined cost function.

The training of artificial neural network is relatively easy but, to a certain extent, they have been sidelined by other machine learning methods in remote sensing applications, such as genetic algorithms and support vector machines. This is mainly because in empiric tests GA and SVM often have performed better than ANN in prediction and classification tasks (Frias-Martinez et al. 2005, Pohjankukka et al. 2018).

3.2.8 Convolutional neural networks

Convolutional neural network (CNN) are a class of deep neural networks, often applied to analyzing visual imagery. CNNs are a form of deep learning, which refers to the use of multiple hidden layers in the network. CNNs typically are fully connected networks, where each neuron in one layer is connected to all neurons in the next layer. They utilize the classifying capability of ‘traditional’ ANN while simultaneously utilizing the recognition of features enabled by the convolutions. As the genetic algorithms, also CNN was inspired by biological processes; whereas GA imitates natural selection, CNN somewhat resembles the generation of visual sensation and shape or pattern recognition by biological organism.

The "fully-connectedness" of these networks makes them also prone to overfitting data. This often necessitates that very large training data sets are used with CNN. CNNs require less pre-processing compared to other image classification algorithms and they can construct complex connections between the elements of input data (e.g. pixels/points), which means that CNN can be trained to learn features that need to be user-provided for traditional (non-deep learning) algorithms. A major advantage of CNN is the independence from prior knowledge of feature design.

In the recent years, the interest in DL utilization has increased, especially due to convolutional neural networks (CNN) having become de-facto method in visual recognition tasks. CNNs have shown their potential in Earth observation (EO), as they have been used to detect over 1.8 billion trees from very high-resolution satellite images in West Africa (Brandt et al 2020), to globally map and date burned areas worldwide (Pinto et al 2020), and to segment floods from radar imagery (Nemni et al 2020). While the most common use cases for CNNs in remote sensing are various classification tasks, such as land use and land cover (LULC) tasks or scene classification tasks, they have also shown potential for predicting continuous targets. For example, CNNs have been used to estimate the amount of cyano-bacteria in river water from aerial hyperspectral images (Pyo et al (2019)). The major downsides of DL are the heavy computational costs and requirement for large, labeled datasets. Thankfully, both of these problems can be in certain amounts addressed with transfer learning (using DL model trained with large, labeled dataset as a feature extractor) or self-supervised learning (first solving simpler, automatically generated task and then applying transfer learning to the original problem), but utilizing these methods with EO data is not yet well studied.

CNN has proved to be very successful in its' original purpose, i.e. in image recognition, which has been a very traditional machine learning task. Also, in the remote sensing of forests the method has shown potential for improving forest estimates and classifications compared to more traditional classifiers and estimators (Ayrey and Hayes 2018, L. Wang et al. 2019, Annala et al. 2020).

3.2.9 Conclusions

The use of existing machine learning algorithms provides a means for better utilizing the full potential of current state of the art remote sensing data in predicting forest characteristics. Based on literature study, several machine learning methods, such as RF, GA or SVM, have produced relatively similar estimation and classification results when using high-dimensional input data sets. Based on these methods it is possible to select a list of features that results in significantly better predictions compared to the use of all potential features.

Generally speaking, no matter which ML algorithm is used, utilizing only optical data is usually not enough for accurate predictions for structural variables such as diameter at breast height (DBH), basal area or growing stock volume. Halme et al (2019) used SVM and Gaussian process regression (GBR) for predicting mean height, basal area, stem biomass, leaf area index (LAI) and main tree species from both aerial hyperspectral images and Sentinel-2 satellite imagery. However, compared to studies utilizing also LiDAR data (e.g. Tuominen et al (2017)), they achieved considerably worse overall results. For example, Halme et al (2019) report 37% relative RMSE for mean height and 45% for basal area at the plot level, whereas Tuominen et al (2017) achieved 10% and 24% respectively for the same forest attributes.

The expected performance of sequential forward/backward selectors is not good in high-dimensional input data, due to their tendency to get stuck in local optima. Thus, they are excluded from further experimental analysis.

Most machine learning methods are still limited by the fact that they require user derived features as their input, and they are not capable of inputting new information to the prediction from the raw remote sensing data.

Deep learning methods such as deep neural networks including CNN are not limited to user-derived features. They can scan through raw remote sensing data (assuming that the data is in regularized form) in order to form their own features for the task underway. On the other hand, with deep neural networks much of that process occurs in a black box, which makes it difficult for the user to extract that new information for further applications.

3.3 Data and methods

3.3.1 Indicators of forest fragmentation from existing spatial data

To develop indicators for forest fragmentation, our aim was to see how well the modeling framework proposed by Estreguil et al. (2014) can be applied for calculating landscape indicators for ecosystem condition accounting in Finland. The framework combines and harmonizes information on habitat pattern, fragmentation and connectivity. In this study, however, connectivity was left out because of difficulties related to defining the required dispersal ability parameter, which is species specific (Estreguil et al. 2014). On the other hand, Czúcz et al. (2018c) pointed out that fragmentation and connectivity are correlated and thus redundant concepts; since the more fragmented a landscape is, the less connectivity it has.

We calculated 15 different forest fragmentation indicators adapting the methodology developed in Estreguil et al. (2011, 2014) and implemented over European forest ecosystems (EEA 2015). These indicators are also part of the Streamlining European Biodiversity Indicators set (SEBI 013, Fragmentation of natural and semi-natural areas). The SEBI was a process which was set up in a response to a request from the EU Environment Council to monitor progress towards reaching the targets of the EU Biodiversity strategy to 2020 (EEA 2012). Given this background, the SEBI indicators offer a solid background for further development of ecosystem condition indicators in Finland.

The European maps and statistics of natural/seminatural landscape and forest fragmentation patterns (EEA 2015) are based on CORINE land cover maps at 1 ha grain size. Both ecological and spatial detail are adequate for mapping these patterns in European scale, but in order to apply the SEBI indicators in national scale, we chose input data produced by Finnish public authorities.

Two main data sources were used in this study. To characterize the general landscape, we used the topographic map database[2] by National Land Survey of Finland, while the MS-NFI data[3] by National Resources Institute Finland (LUKE) was used to characterize different focal boreal forest habitats (i.e. ecosystem types). Both datasets were identified as key datasets for ecosystem accounting in an earlier study by Vihervaara et al. (2018) and are publicly available in spatial format. MS-NFI data is available at 20 m grain size for 2009, and at 16 m for years 2011, 2013, 2015 and 2017. The topographic map database is in vector format at 1:10 000 nominal scale and being constantly updated and thus only the latest version is publicly available. However, since the topographic map database is used in the production of MS-NFI as one of the input data layers to characterize all other LULC types other than forest, we were able to retrieve time-series of this data from LUKE's archives. Using the topographic data together with MS-NFI

---

[2] https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/topographic-database

[3] http://kartta.luke.fi/index-en.html

was straightforward as the topographic data was already rasterized and aligned to the same resolution and map grid in the MS-NFI production process (Mäkisara et al. 2019).

An important note is that before doing any further processing, both MS-NFI and topographic map data from 2009 had to be first resampled from their native resolution of 20 m to 16 m to allow comparisons with later years. In concrete terms, it means that the baseline data of 2009 has less detailed information than the data it is compared to, which is good to keep in mind when interpreting the results.

Since all SEBI-013 indicators are based on area proportions of the total landscape, the methodology can be applied at any scale – from local (e.g. municipalities, urban areas) to regional (e.g. administrative regions, watersheds) and national (e.g. whole Finland), as well as for any ecosystem type where suitable spatial data exists. In SEEA-EEA terminology, the scale of application would be equivalent to the ecosystem accounting area (EAA). To demonstrate the calculation of these indicators, we avoided the unnecessary computational burden of doing the whole country, but instead chose one MS-NFI mapsheet (M4) in Southern Finland, covering total area of 18 432 km^2 and supporting a mosaic of different land use and forest types. As for the focal habitat, we concentrated on three different forest types based on tree species dominance: pine, spruce, and deciduous trees, and a fourth focal habitat type which consisted of all forests combined[4]. Lastly, to be able to see if any change happened in the indicator values and to set reference levels for each indicator, we calculated the indicators for two accounting years: 2009 and 2017. The indicator values of 2009 were also chosen as the reference level values.

Based on Estreguil et al. (2014), we used three different but complimentary landscape modelling approaches, each giving a different perspective on patterns of natural and seminatural areas and forest fragmentation:

1. **General landscape model**. The landscape composition is summarized by measures describing the focal habitat in a landscape in terms of its total amount, its pattern and landscape context.
2. **Morphological pattern model**. The focal habitat cover is divided based on their morphology into core areas, boundaries, linear features and islets. Fragmentation of the focal habitat type is characterized by indicators quantifying the share of focal habitat area in these morphological classes.
3. **Edge interface model**. This model is derived from the integration of the two first models. Focal habitat edge interfaces are differentiated by morphology (edges or boundaries, linear features and islets) and characterized according to the similarity of adjacent habitats (whether the edge is situated next to natural/seminatural or anthropogenic/artificial habitats). In other words, the indices derived from the edge interface model give us a measure of edge naturalness.

A critical parameter for these models is the **edge width**. The edge width is set by user for the morphological spatial pattern analysis (MSPA), as it determines the width of the focal habitat boundary, and subsequently impacts the patterns and proportions of interior, linear and islet landscape elements, and any indicators derived from them. The same edge width should also be used for the disk radius of the neighborhood, when applying the general landscape model (see next chapter). Selecting the appropriate edge width is thus a fundamental assumption for these models, but not an easy one as different organisms and species respond differently to edges (Öhman & Eriksson 1998). In forestry the edge width is related to the height and structure of the forest, ranging from narrow (20 m) to wide (160 m) (Franklin & Forman

[4] Altogether more than 30 tree species can be found in Finnish forests, out of which the most common are Norway spruce, Scots pine, Silver birch and Downy birch. All 30+ species are found in the sample plots of NFI, but in the MS-NFI, species-level data is available only for Norway spruce. Scots pine and all conifers other than Norway spruce form the group “pine” whereas Silver and Downy birch are handled as a single “birch” species group. Apart from these four most common species, the remaining species (most of them deciduous) are economically insignificant and very rarely form homogeneous stands and are thus pooled into species group “other deciduous trees”. For practical reasons, we decided to group the birch species group and other deciduous trees together.

1987). Estreguil et al. (2014) used edge width of 25 m, and for boreal forests edge widths of 32 and 64 m were suggested in an earlier study by Öhman & Eriksson (1998). For this study, we experimented with three edge widths (48, 112 and 240 meters), out of which 48 meters gave most feasible results.

The following three sub-chapters give additional detail on the methodology used for indicator calculation. All GIS processing steps followed the methodology outlined in Estreguil et al. (2011, 2014) with some adaptations to Finnish context. Only free and open-source software, namely QGIS (general raster processing, visualization), GRASS (complex raster operations) and Guidos Toolbox (morphological spatial pattern analysis) were used to perform all analysis steps, which are documented either in graphical models or processing scripts enabling easy replication and revision.

3.3.1.1 General landscape indicators

Habitat pattern in the landscape affects the interactions between and within species both within and between patches (Estreguil et al. 2014). The landscape context of habitats influences habitat content (e.g. forest condition), and thus offers feasible indicators for assessing ecosystem condition.

First, to describe the general landscape composition, we reclassified the 38 LULC classes (see table 2.4 in Mäkisara et al. 2019 for details) of the rasterized topographic map (Figure 13a) into three separate binary raster maps (Figure 13b):

- MS-NFI classes 1-20 (built-up features, urban green areas other than forests, infrastructure, quarries and peat production areas) as **artificial**
- MS-NFI classes 21-22 (meadows and agricultural fields) as **agriculture**
- MS-NFI classes 91-112 (water, bare sand and bedrock, open bogs, paludified areas, forests and forested marshes) as **natural**

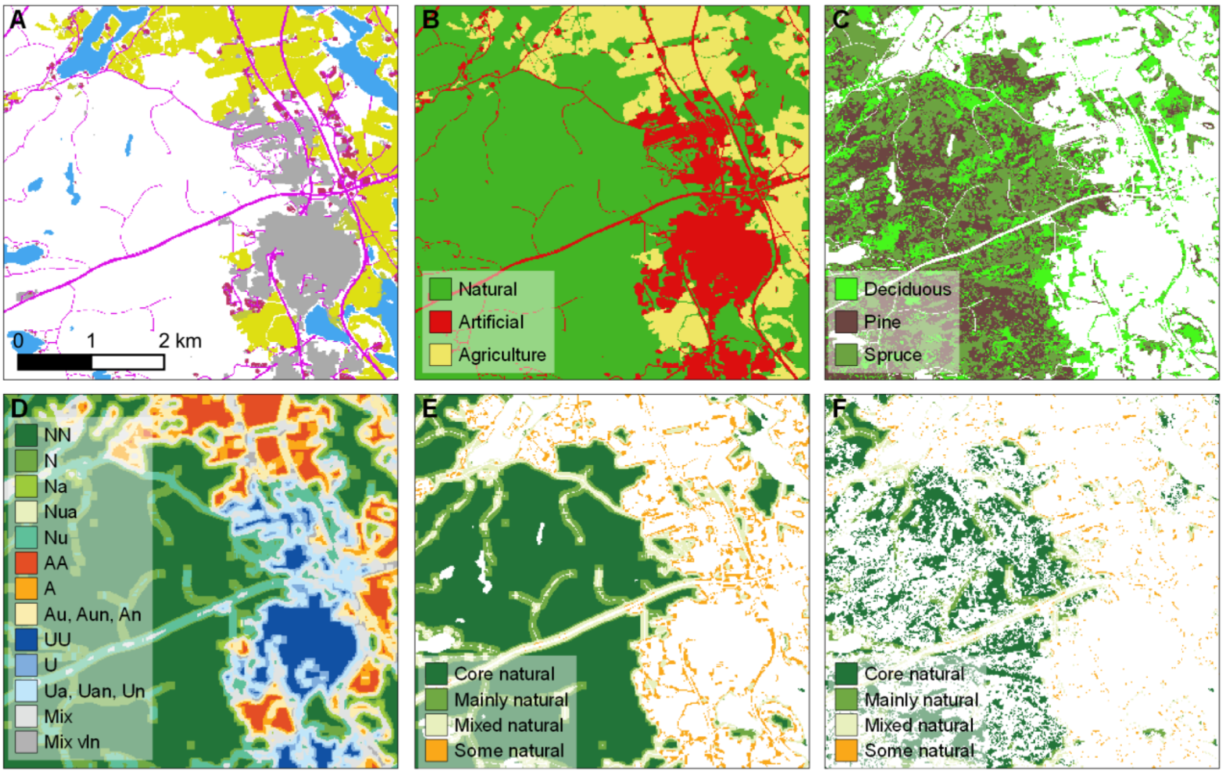

Figure 13: Processing steps for landscape mosaic: Rasterized topographic map (A), reclassified into three general landscape components (B), focal habitats (C), landscape mosaic index (D), simplified landscape mosaic extracted for a focal habitat: all forests (E) and spruce (F). For landscape pattern type explanations, see figure 15.

Second, we extracted total volume[5] of pine, spruce, birch and other deciduous trees as raster maps from the MS-NFI data. With raster calculator expressions, we evaluated the dominant tree species for each pixel, with birch and other deciduous trees treated together. This resulted in four binary raster maps describing the patterns of the focal habitat, where the focal habitat could be either pine-, spruce-, or deciduous dominated, or all forest types put together (Figure 13c). These focal habitat maps can also be considered as a data source from which **forest ecosystem extent indicators** can be complied (Table 23).

Table 23: Forest ecosystem extent indicators following draft SEEA-EA guidelines (UNSD 2020: table 14.1)

Indicator	Measurement unit
Percentage of EAA covered by specific forest habitat type	Hectares, % of opening
Change of area covered by specific forest habitat types during an accounting period	%
Percentage of area covered by specific forest habitat type that is unchanged (opening stock – reduction)	Hectares; % of opening
Percentage of area covered by specific forest habitat type which has changed (additions + reductions)	Hectares; % of opening

Third, we applied a moving window approach with neighborhood size of 7 pixels to obtain the landscape mosaic context of each pixel in the EAA. In other words, for every pixel we counted how many pixels of artificial, agriculture and natural landscape exists in the neighborhood of 7 x 7 = 49 pixels (or 1.3 ha). The pixel count was then converted to proportions, which resulted in three raster maps storing the areal proportion (0 – 100 %) of each class for each pixel.

Fourth, we used raster calculator to classify each pixel in the EAA into one of seventeen landscape pattern types based on pre-defined thresholds (Figure 14) and proportion of artificial, agriculture and natural[6]

[5] The volume of a tree in MS-NFI is defined as the volume of the stem wood above stump until the top of the tree.

[6] A word of caution: The term ‘natural’ as used in the modelling framework presented in this study does not refer to naturalness as a condition in ecological sense, but simply refers to the landscape context in a strict and straightforward division of the landscape into three components, natural, agricultural, and urban. The natural here also includes semi-natural landscapes. Naturalness of a forest habitat cannot be determined from MS-NFI data alone. The great majority of Finnish forests are managed and forest habitats which could be considered to be in natural state are very rarely found outside protected areas (but managed forests are also found inside protected areas!)

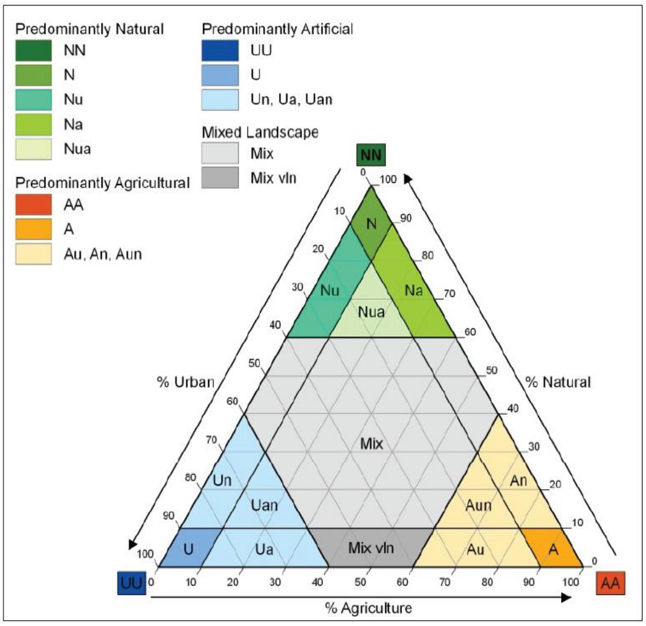

Figure 14: The 17 landscape pattern types derived from the landscape mosaic index (from Estreguil et al. 2011)

calculated in the previous step. The output of this operation is a new raster map called *landscape mosaic index* (Figure 13d) (Wickham & Norton 1994, Riitters et al. 2000, Riitters et al. 2009).

These seventeen classes were simplified into four classes of interest following Estreguil et al. (2014):

- Focal habitats in **core natural** patterns (NN) are always (100 %) adjacent to natural/seminatural habitats or in the interior part of patches,
- Focal habitats in **mainly natural** patterns (N) are mainly (80%) bordering natural/semi-natural habitats,
- Focal habitats in **mixed natural** patterns (MN = Na + Nu + Nua) are embedded in a predominant natural context (Nx), but are significantly fragmented by agricultural and/or artificial land,
- Focal habitats in **some natural** patterns (SN = sum of all the other types) are predominantly in human induced contexts (Ux, Ax, Mix) (i.e. forest patch in agricultural landscape).

Finally, we intersected the simplified landscape mosaic with each four focal habitats (Figure 13e and Figure 13f). A total of six indicators describing the general landscape context can calculated (Table 24). The first indicator, proportion of all natural/seminatural habitat in landscape (NP), is a general indicator and cannot be calculated for a focal habitat. The other five indicators in this group can be calculated for any given focal habitat type and describe its share of the total landscape (FP) as well as in different landscape mosaic contexts (NN-P, N-P, MN-P and SN-P).

Table 24: Forest ecosystem condition indicators characterizing the general landscape content

Indicator	Abbreviation	Measurement unit
Proportion of natural/semi-natural habitat in landscape	NP	%
Proportion of the focal habitat in landscape	FP	%
Proportion of focal habitat in 'core natural' landscape mosaic context	NN-P	%
Proportion of focal habitat in 'mainly natural' landscape mosaic context	N-P	%
Proportion of focal habitat in 'mixed natural' landscape mosaic context	MN-P	%
Proportion of focal habitat in 'some natural' landscape mosaic context	SN-P	%

3.3.1.2 Morphological pattern indicators

Morphological shapes of forest habitat, such as amount and connectivity of interior areas, geometry of habitat edges, and presence of clumps and linear strips of habitat in the landscape all play an important ecological role in forest ecosystem condition (Estreguil et al. 2014).

The mathematical morphological spatial pattern analysis (MSPA) application Guidos Toolbox, developed by Soille and Vogt (2009), was used to implement the model. The only input data required is a spatially referenced raster maps of the focal habitat in the landscape. We used the same focal habitat maps as described earlier, but as an additional processing step, the maps had to follow the modelling software requirements. This required combining the focal habitat maps with all forests map and recoding pixel values as following (see Figure 15a and Figure 15c):

- ignored, includes all forests *excluding* the focal habitat[7]
- background, non-forest (everything else)
- foreground, focal habitat (pine/spruce/deciduous/all forests)

Guidos Toolbox segments the input raster map and outputs twenty-five classes which are further simplified into six mutually exclusive spatial pattern classes (see Figure 15b and Figure 15d): a) **interior** or **core** areas which are beyond a fixed distance to the border (defined by *edge width*), b) **boundaries** of interior areas either on the outer side (*edges*) or internal side (*perforations*), c) **connectors** which consist of focal habitat with no core and connects either to two or more different core habitats or to the same core (*loops*), d) **branches** which consist of focal habitat with no core and are connected to interior areas from only one end, e) physically isolated **islets**, too small to form an interior, and f) background, which includes everything else.

[7] In the case when focal habitat is all forests, value 0 is not used

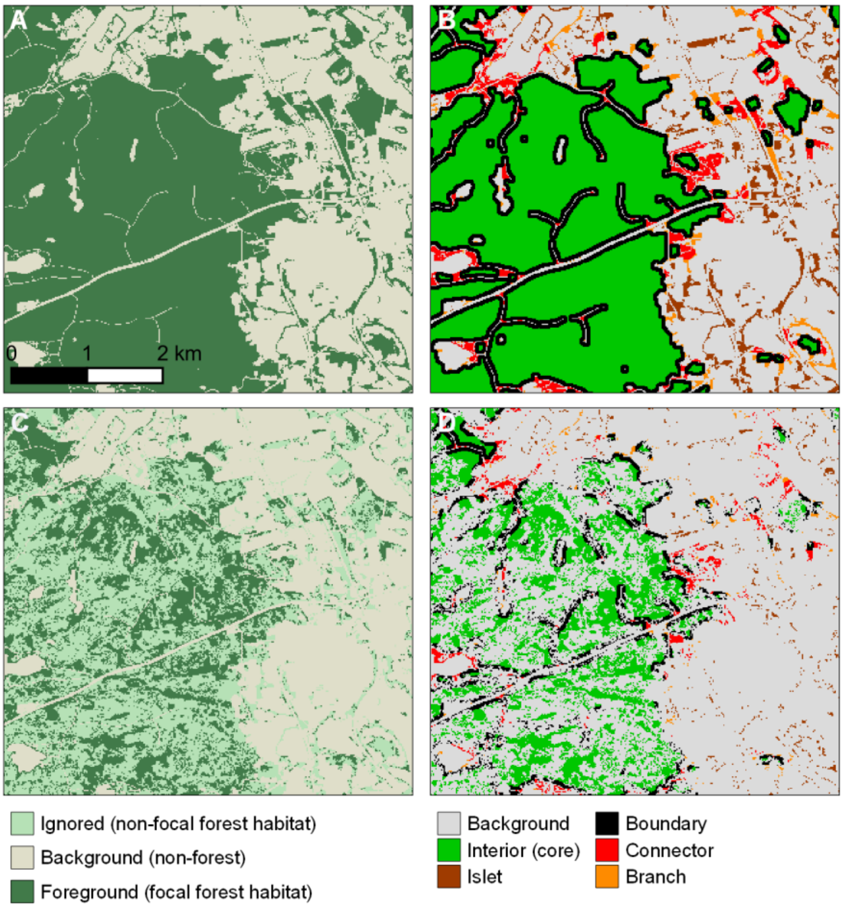

Figure 15: Processing steps for morphological pattern model: input forest mask for all forests (A), simplified model output for all forests (B), input forest mask for spruce (C) and simplified model output for spruce (D)

Four indicators describing morphological patterns can be calculated for each focal habitat (Table 25). These indicators are simple quantifications of the proportions of a focal habitat in morphological classes: in the interior (IFP), in islets (IS), and in linear features (LIP) which includes boundaries, connectors and branches. A special case is IF*P, which is IFP enlarged with by the NN (natural) part of the boundary class (see next chapter).

Table 25: Forest ecosystem indicators characterizing morphological patterns

Indicator	Abbreviation	Measurement unit
Proportion of focal habitat in interior	IFP	%
Proportion of focal habitat in interior and in boundaries with natural edge interface (IF + SI-BO$_{NN}$)	IF*P	%
Proportion of focal habitat in islet	ISP	%
Proportion of focal habitat in linear features (boundary, connector and branch)	LIP	%

3.3.1.3 Edge interface indicators

The third landscape modeling component is simply a combination of the general landscape and morphological pattern maps. Following the conceptualization of Estreguil et al. (2014), focal habitats in *core natural* (NN) and *mainly natural* (N) patterns are considered to have ‘soft’ edge interface context classified as **natural**, whereas focal habitats in mixed natural (MN) and some natural (SN) have ‘hard’ edge interfaces classified as **artificial**.

Figure 16 illustrates the processing steps for the edge interface model. For each focal habitat, the landscape mosaic model is intersected with the morphological model to discriminate between the interior and edge part of habitat patches, and among different edge morphologies (boundary, linear feature and islet). Following Estreguil et al. (2014), six similarity indicators (SI) are derived from the edge interface model (Table 26). These indicators quantify the proportion of a given focal habitat in boundaries, islets and linear features with either natural or artificial edge interfaces.

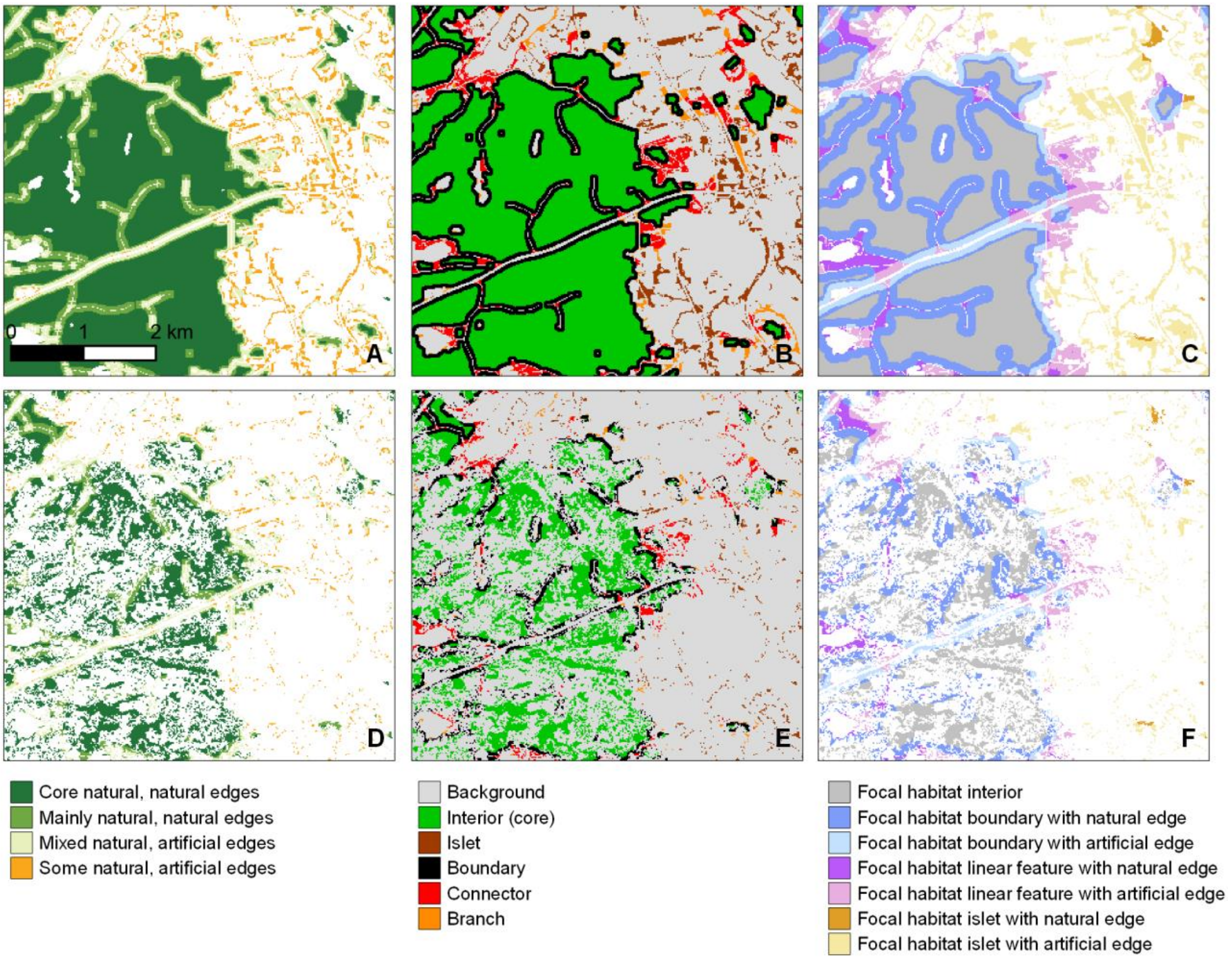

Figure 16: Processing steps for edge interface model: the landscape mosaic model (left column) and the morphological model (middle column) are intersected to get the edge interface model (right column). Examples shown for all forest (A, B, C) and spruce (D, E, F) focal habitats

Table 26: Forest ecosystem condition indicators characterizing edge interface.

Indicator	Abbreviation	Measurement unit
Proportion of focal habitat in boundary with natural edge interface	$SI\text{-}BO_{NN}$	%
Proportion of focal habitat in boundary with artificial edge interface	$SI\text{-}BO_{OTHER}$	%
Proportion of focal habitat in islet with natural edge interface	$SI\text{-}IS_{NN}$	%
Proportion of focal habitat in islet with artificial edge interface	$SI\text{-}IS_{OTHER}$	%
Proportion of focal habitat in linear feature with natural edge interface	$SI\text{-}LI_{NN}$	%
Proportion of focal habitat in linear feature with artificial edge interface	$SI\text{-}LI_{OTHER}$	%

3.3.2 Ecosystem indicators from remote sensing data

Our ground reference data consisted of around 1500 field plot with 9 m radius over an area of ca. 5 800 km² and includes total and species-wise growing stock (m³/ha), mean diameter at breast height (cm) as well as mean height for each plot (Table 27). As the remote sensing data, we used aerial false-color images with ground resolution of 0.3 m, as well as LiDAR data with average point density of 1.66 pts/m². The plots were split to training, validation and test set with 70-15-15 -split. The locations of the field plots are presented in Figure 17.

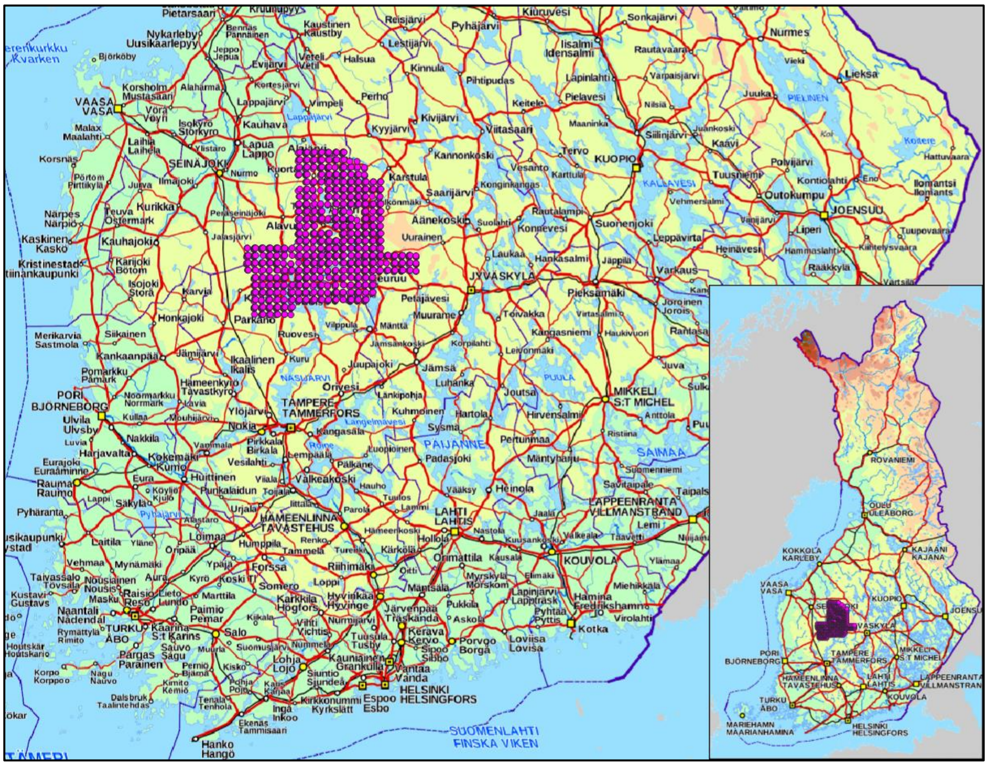

Figure 17: Locations of ground reference plots in central Finland.

The most common tree species in our field plots was pine, both in the sense of average species-wise volume and average proportion of the species. Pine was the dominant tree species group in around 73% of the tree species, whereas spruce and deciduous trees were the dominant species in 16% and 11% of the plots respectively. Overall, the field plots represented a wide variety of forest structures and species group differences and were representative for the corresponding geographical area.

Table 27: Summary of forest stand variables in training data

	Volume (m^3/ha)	Height (m)	DBH (cm)	Pine (m^3/ha)	Spruce (m^3/ha)	Deciduous (m^3/ha)
Mean	145.71	14.65	17.89	89.05	34.34	22.33
Std	86.10	4.14	5.39	71.01	66.26	35.75
Min	2.70	3.80	5.50	0.00	0.00	0.00
Max	867.35	28.10	36.20	519.86	424.98	288.48

We used the k-nearest neighbors with genetic algorithm (KNN-GA) (Tuominen et al. 2013) as the reference method, as similar method is used to produce MS-NFI in Finland, and compared the performance of RF, ANN and CNN with it. As the CNN architectures, we used the architecture presented by Ayrey et al (2018) for LiDAR data and DenseNet by Huang et al (2016) for aerial images. As combining LiDAR and optical data for CNNs proved to be difficult, we used a two-step process for species-wise volume prediction with CNNs. First, we transformed all species-wise volumes into proportions for each field plot, and then trained a model to predict those from aerial images. The predicted proportions were then used to get the species-wise volumes based on predicted total volume. In order to ensure that predictions for species-wise proportions sum to 1, we included a normalization layer as a final layer for the DenseNet implementation, where each of the predicted values was divided by the sum of all predictions for the corresponding field plot.

Even though CNNs are able to process raw image data, the same is not true with LiDAR point clouds, as CNN expect the data to have a grid-like topography. Because of this, the LiDAR point clouds were *voxelized* into 45x45x40cm voxels (see Figure 18 for visualization), and the value for each voxel was assigned to be 1 if any points were within the voxel and 0 otherwise. We also tested whether using the number of returns within each voxel as the values affected the performance, but it had practically no effect. For other methods, we extracted both optical and textural features from the false-color images, and a set of point cloud metrics, such as proportions of points in various height thresholds, from LiDAR point clouds (Tuominen et al. 2020).

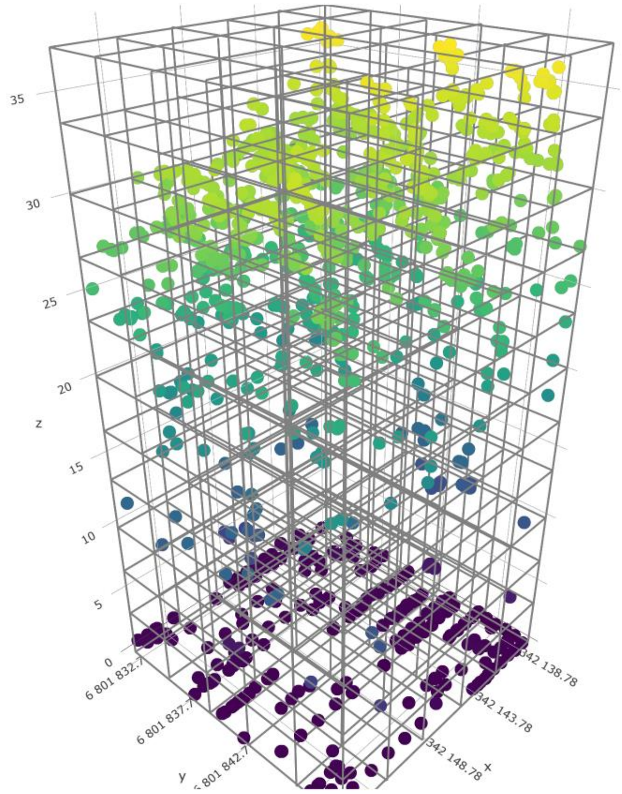

Figure 18: Visualization of voxelized field plot. Each voxel represents the points within the corresponding 3D space.

We used root mean squared error (RMSE), relative RMSE (RMSE-%), bias and relative bias (bias-%) to evaluate our models. RMSE-% and bias-% were acquired by dividing RMSEs and biases with the mean value of the targets and multiplying them by 100. For neural network models (ANN and CNN), we used mean squared error as the loss function to optimize the models.

3.4 Results and discussion

3.4.1 Indicators of forest fragmentation

Indicators of forest fragmentation were calculated for a whole MS-NFI mapsheet, but for the sake of demonstration, results for two municipalities with very contrasting characteristics are shown here (Figure 19). Akaa is a small town of 16 000 inhabitants with two bigger urban centres. Agriculture forms major part of the landscape and roughly one-third of the landscape area is forested. Padasjoki, on the other hand, has relatively small amount of agriculture and artificial areas in the landscape. Forests are occupying nearly half of the landscape which is otherwise dominated by other natural areas, mostly lakes.

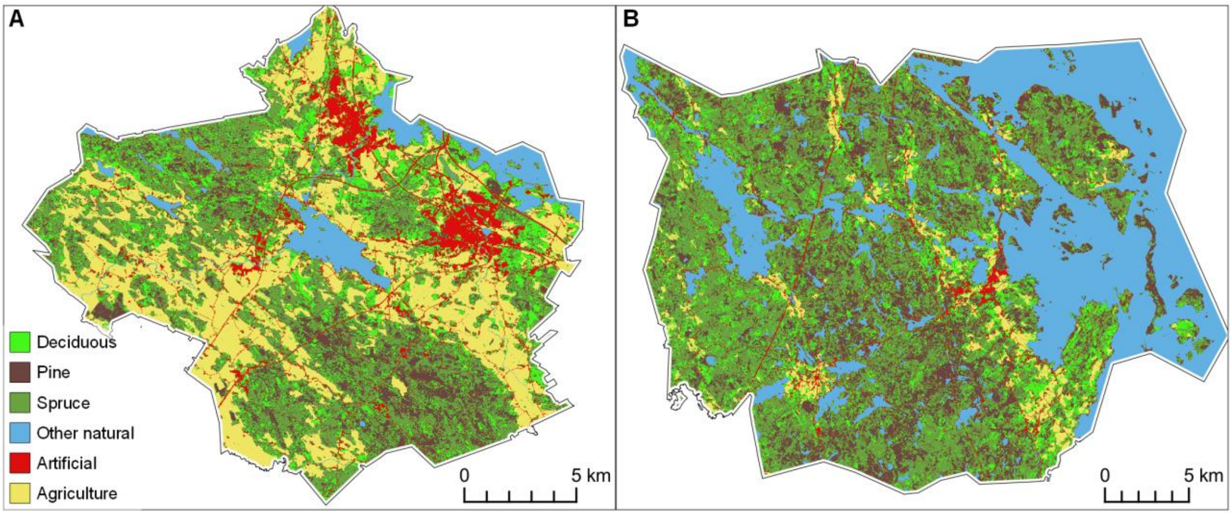

Figure 19. Focal habitat and general landscape patterns (2017) of two municipalities used for demonstration: Akaa (A) and Padasjoki (B).

Demonstrations of ecosystem condition indicator accounts for these municipalities are shown in Table 28 and Table 29. The first indicator, NP, verifies the visual interpretations that can be made from Figure 19: the proportion of natural and seminatural habitats in Padasjoki (0.681) is twice as much as in Akaa (0.347), and has remained practically unchanged between 2009 and 2017. Proportion of forest habitats in the landscape (FP) follow the typical order in Finland: spruce-dominated forest type is the most common, followed by pine and deciduous species. An important observation is that deciduous forests are more often found outside core natural landscape mosaic context than pine and spruce (NN-P, N-P, MN-P, SN-P), which indicates a stronger fragmentation pattern. We must stress, however, that the analysis is done purely from landscape perspective, and we cannot say what are the *causes* behind these fragmentation patterns. As stated earlier, these indicators describe forest *state*, not the *pressures* or *processes* that cause the fragmentation patterns observer here. We can only speculate how much the fragmentation is caused by pressures of anthropogenic origin and how much can be explained by natural ecosystem processes, as deciduous trees in boreal forests naturally tend to grow in mixed stands with coniferous trees.

Ratio of interior versus non-interior habitat is an important measure of fragmentation (Estreguil et al. 2014). When looking at the morphological patterns in the example, we get further hints of fragmentation of deciduous forests. The proportion of forest habitat in interior (IFP) as well as in interior and in boundaries with natural edge interface together (IF*P) is considerably smaller for deciduous than for pine and spruce forests. This trend is stronger in Akaa than in Padasjoki. This naturally leads to a situation where a bigger proportion of focal habitat situates in islets (ISP) and in linear features (LIP). While islets offer important habitat provision services, physical isolation and lack of connectivity to other forest patches can worsen their condition and makes them vulnerable to disappear due to their shape and size (Estreguil et al. 2014). Linear features, on the other hand, also provide critical ecosystem services – especially connectors, which provide ecological corridors critical for species movement and dispersal – but are as well endangered by fragmentation and other anthropogenic pressures.

Fragmentation also relates to the shift in land use at edges: permeability of natural edge interfaces for species dispersal is higher than those of artificial edges (Estreguil et al. 2014). As expected, we see evidence of higher level of fragmentation for deciduous forests than the other two forest types: whereas proportions of focal habitat in boundaries, islets and linear features with natural edge interface ($SI\text{-}BO_{NN}$, $SI\text{-}IS_{NN}$, $SI\text{-}LI_{NN}$) are more even between the three forest types, the same proportions with artificial edge interface ($SI\text{-}BO_{OTHER}$, $SI\text{-}IS_{OTHER}$, $SI\text{-}LI_{OTHER}$) are considerably larger for deciduous than other forest types, and more so in Akaa than in Padasjoki.

3.4.1.1 Uncertainties

Since the forest fragmentation indicators are derived from existing national datasets on forest inventories and land use, it is clear that any results derived by analyzing this data will inherit the uncertainties and errors present in the source data. Assessing uncertainties and errors becomes even more complex as the source data is a time series. Pixel-level prediction errors in the MS-NFI are generally rather high (Tomppo et al. 2013) but decrease when the area in question increases. The errors vary by predicted forest variable and depend also on the actual value in the field, for example on the volume of growing stock and the site fertility class. Several error sources have been identified in Tomppo et al. (2008). However, the densities of field plots used in the inventory are high enough to ensure that the resulting sampling errors are low for core variables, such as volume of growing stock (Mäkisara et al. 2019).

Availability of suitable satellite images for MS-NFI has been one of the biggest sources of uncertainty (Tomppo et al. 2013). For example, in the 2009 inventory, only two satellite sensors (Landsat-5 and 7) with suitable specifications were available for analysis (some coarser resolution satellites were also used as alternative). The combination of short vegetation growth period, 16-day repeat cycle and frequent cloud cover, typical situation in Nordic summer, further complicated wall-to-wall mapping of forest parameters. In fact, the 2009 data still has some number of no-data pixels where no estimations could be made. The situation changed considerably in the MS-NFI of 2017, as Landsat-5 was decommissioned and data from new satellite sensors with higher radiometric resolution (Landsat-8) and higher spectral as well as spatial resolution (Sentinel-2) became available. Not only did the quality of data improve, but also the amount of available satellite images, improving the probabilities of collecting data with clear-sky conditions.

The forest extent and fragmentation indicators presented in this study depend in one MS-NFI forest parameter, the volume estimates. For the 2009 inventory[8], the magnitude of average errors of volume estimates at pixel level in Southern Finland for all species were 86 and 62 m^3/ha in mineral and peat soil (Tomppo et al. 2008). Errors of volume estimates disaggregated by species group were largest for pine (63 and 50 m^3/ha in mineral and peat soil) and smallest for other deciduous trees group (22 and 9 m^3/ha). However, when the MS-NFI volume estimates are aggregated to municipal, regional or national scale, the results are more robust as pixel-level errors are smoothed in larger scale.

Until 2011, the MS-NFI pixel size was 20 m, which was changed to 16 m in 2013. In this study, the 2009 MS-NFI data was resampled to match the 16 m pixel size and to make the results comparable. Changing the spatial resolution can affect raster values at pixel level, but at larger scales this effect is smoothed out.

The topographic map database has positional accuracy comparable to maps on scale 1:5 000–1:10 000 and consists of map sheets which are updated in 5–10 year periods, but some elements such as roads and buildings are updated yearly (Mäkisara et al. 2019), which can lead to temporal inconsistencies in the data.

[8] Error estimates are not available for the 2017 MS-NFI.

Table 28: Demonstration of ecosystem condition indicator account for Akaa municipality using landscape level characteristics. Indicators are given for the reference condition (2009), accounting year 2017, and absolute change in indicator values between the accounting year and reference condition.

Ecosystem Condtion Typology: <u>Landscape level charateristics</u>			Accounting area		Akaa		Accounting area		Akaa		Accounting area		Akaa	
			Reference year		**2009 (reference)**		**Accounting year**		**2017**		**Accounting year**		**CHANGE**	
			Extent (ha)		56491		Extent (ha)		56491		Extent (ha)		56491	
			Edge width (m)		48		Edge width (m)		48		Edge width (m)		48	
			Neighborhood		7 x 7		Neighborhood		7 x 7		Neighborhood		7 x 7	
			Ecosystem type (i.e. focal habitat)				Ecosystem type (i.e. focal habitat)				Ecosystem type (i.e. focal habitat)			
	Description	Indicator	All forests	Deciduous	Pine	Spruce	All forests	Deciduous	Pine	Spruce	All forests	Deciduous	Pine	Spruce
General landscape	Proportion of natural/semi-natural habitat in landscape	NP	0,347				0,347				0,000			
	Cover of the focal habitat in landscape (ha)	FC	17519	4069	4711	8568	17512	5433	4846	7519	-7	1364	135	-1049
	Proportion of the focal habitat in landscape	FP	0,310	0,072	0,083	0,152	0,310	0,096	0,086	0,133	0,000	0,024	0,002	-0,019
	Proportion of focal habitat in 'core natural' landscape mosaic context	NN-P	0,565	0,432	0,596	0,610	0,560	0,455	0,642	0,580	-0,005	0,023	0,045	-0,029
	Proportion of focal habitat in 'mainly natural' landscape mosaic context	N-P	0,159	0,154	0,149	0,166	0,165	0,169	0,145	0,174	0,006	0,015	-0,004	0,008
	Proportion of focal habitat in 'mixed natural' landscape mosaic context	MN-P	0,180	0,230	0,165	0,166	0,180	0,213	0,137	0,184	0,000	-0,016	-0,028	0,018
	Proportion of focal habitat in 'some natural' landscape mosaic context	SN-P	0,096	0,184	0,089	0,058	0,096	0,163	0,076	0,062	0,000	-0,022	-0,013	0,004
Morphology	Proportion of focal habitat in interior	IFP	0,572	0,440	0,606	0,619	0,573	0,463	0,655	0,593	0,000	0,023	0,049	-0,027
	Proportion of focal habitat in interior and in boundaries with natural edge interface (IF + BO_{NN})	IF*P	0,710	0,563	0,730	0,760	0,711	0,601	0,776	0,737	0,001	0,037	0,046	-0,023
	Proportion of focal habitat in islet	ISP	0,047	0,112	0,056	0,032	0,048	0,090	0,038	0,031	0,000	-0,021	-0,018	0,000
	Proportion of focal habitat in linear features (boundary, connector and branch)	LIP	0,380	0,449	0,338	0,349	0,380	0,447	0,307	0,376	0,000	-0,002	-0,031	0,027
Interface	Proportion of focal habitat in boundary with natural edge interface	$SI\text{-}BO_{NN}$	0,137	0,124	0,125	0,141	0,138	0,138	0,122	0,145	0,001	0,014	-0,003	0,004
	Proportion of focal habitat in boundary with artificial edge interface	$SI\text{-}BO_{OTHER}$	0,136	0,136	0,107	0,126	0,138	0,140	0,100	0,140	0,002	0,005	-0,007	0,014
	Proportion of focal habitat in islet with natural edge interface	$SI\text{-}IS_{NN}$	0,002	0,002	0,003	0,002	0,002	0,002	0,001	0,002	0,000	0,000	-0,001	0,000
	Proportion of focal habitat in islet with artificial edge interface	$SI\text{-}IS_{OTHER}$	0,046	0,109	0,053	0,029	0,046	0,088	0,036	0,029	0,000	-0,021	-0,017	0,000
	Proportion of focal habitat in linear feature with natural edge interface	$SI\text{-}LI_{NN}$	0,012	0,021	0,013	0,013	0,012	0,022	0,010	0,015	0,000	0,001	-0,003	0,002
	Proportion of focal habitat in linear feature with artificial edge interface	$SI\text{-}LI_{OTHER}$	0,094	0,169	0,094	0,070	0,091	0,147	0,076	0,076	-0,003	-0,022	-0,018	0,007

Table 29: Demonstration of ecosystem condition indicator account for Padasjoki municipality using landscape level characteristics. Indicators are given for the reference condition (2009), accounting year 2017, and absolute change in indicator values between the accounting year and reference condition

			Reference year		2009		Accounting year		2017		Accounting year		CHANGE	
			Extent (ha)		98818		Extent (ha)		98818		Extent (ha)		1843200	
			Edge width (m)		48		Edge width (m)		48		Edge width (m)		48	
			Neighborhood		7 x 7		Neighborhood		7 x 7		Neighborhood		7 x 7	
			Ecosystem type (i.e. focal habitat)				Ecosystem type (i.e. focal habitat)				Ecosystem type (i.e. focal habitat)			
	Description	Indicator	All forests	Deciduous	Pine	Spruce	All forests	Deciduous	Pine	Spruce	All forests	Deciduous	Pine	Spruce
General landscape	Proportion of natural/semi-natural habitat in landscape	NP	0,681				0,682				-0,001			
	Cover of the focal habitat in landscape (ha)	FC	46597	7984	13508	24897	46749	9919	15010	23005	152	1935	1502	-1892
	Proportion of the focal habitat in landscape	FP	0,472	0,081	0,137	0,252	0,473	0,100	0,152	0,233	0,002	0,020	0,015	-0,019
	Proportion of focal habitat in 'core natural' landscape mosaic context	NN-P	0,665	0,545	0,683	0,693	0,655	0,544	0,679	0,688	-0,010	0,000	-0,004	-0,006
	Proportion of focal habitat in 'mainly natural' landscape mosaic context	N-P	0,181	0,196	0,176	0,178	0,190	0,205	0,179	0,190	0,009	0,009	0,002	0,012
	Proportion of focal habitat in 'mixed natural' landscape mosaic context	MN-P	0,133	0,200	0,123	0,117	0,135	0,194	0,130	0,112	0,002	-0,006	0,006	-0,005
	Proportion of focal habitat in 'some natural' landscape mosaic context	SN-P	0,022	0,060	0,017	0,012	0,021	0,057	0,013	0,011	-0,001	-0,003	-0,004	-0,001
Morphology	Proportion of focal habitat in interior	IFP	0,627	0,534	0,649	0,645	0,624	0,535	0,624	0,664	-0,003	0,002	-0,025	0,019
	Proportion of focal habitat in interior and in boundaries with natural edge interface (IF + BO_{NN})	IF*P	0,817	0,704	0,832	0,837	0,819	0,716	0,815	0,852	0,001	0,012	-0,017	0,015
	Proportion of focal habitat in islet	ISP	0,007	0,022	0,007	0,007	0,007	0,017	0,008	0,005	0,000	-0,005	0,001	-0,001
	Proportion of focal habitat in linear features (boundary, connector and branch)	LIP	0,366	0,445	0,343	0,348	0,369	0,447	0,367	0,331	0,003	0,003	0,024	-0,017
Interface	Proportion of focal habitat in boundary with natural edge interface	$SI\text{-}BO_{NN}$	0,190	0,170	0,182	0,192	0,194	0,180	0,190	0,188	0,004	0,011	0,008	-0,004
	Proportion of focal habitat in boundary with artificial edge interface	$SI\text{-}BO_{OTHER}$	0,086	0,106	0,076	0,074	0,087	0,108	0,069	0,075	0,001	0,002	-0,007	0,001
	Proportion of focal habitat in islet with natural edge interface	$SI\text{-}IS_{NN}$	0,002	0,004	0,002	0,004	0,002	0,002	0,004	0,002	0,000	-0,002	0,002	-0,001
	Proportion of focal habitat in islet with artificial edge interface	$SI\text{-}IS_{OTHER}$	0,005	0,018	0,005	0,003	0,005	0,015	0,005	0,003	0,000	-0,003	-0,001	0,000
	Proportion of focal habitat in linear feature with natural edge interface	$SI\text{-}LI_{NN}$	0,026	0,033	0,026	0,031	0,024	0,032	0,039	0,023	-0,002	-0,002	0,013	-0,007
	Proportion of focal habitat in linear feature with artificial edge interface	$SI\text{-}LI_{OTHER}$	0,064	0,136	0,059	0,051	0,064	0,128	0,069	0,045	0,000	-0,008	0,010	-0,007

3.4.2 Predicting forest stand variables from remote sensing data

For non-species related forest variables, our CNN was able to outperform other methods by using only voxelized LiDAR data, whereas other methods utilized features from both LiDAR point clouds and aerial imagery. The most significant improvement was with total volume, where the RMSE-% improved by around 3 percentage points compared to KNN-GA and ANN, and 5 percentage points compared to RF (Table 30). For both mean height and mean DBH the differences were not as significant, as the difference between best and worst results was less than one percentage point for DBH and around 1.5 percentage points for mean height. Interestingly, for mean height all other methods were able to achieve slightly better results than benchmark method KNN-GA.

Table 30: Results for overall forest variables

	KNN-GA	**RF**	**ANN**	**CNN**
RMSE				
Mean height (m)	1.18	1.06	1.09	0.96
Total volume (m^3/ha)	36.21	39.54	36.59	31.97
Mean DBH (cm)	2.24	2.28	2.25	2.12
RMSE-%				
Mean height (m)	8.04%	7.19%	7.45%	6.55%
Total volume (m^3/ha)	23.82%	26.01%	24.07%	21.03%
Mean DBH (cm)	12.42%	12.64%	12.49%	11.73%
BIAS				
Mean height (m)	-0.12	0.02	-0.08	0.14
Total volume (m^3/ha)	3.21	3.71	-1.16	-3.70
Mean DBH (cm)	-0.18	0.09	-0.06	-0.06
BIAS-%				
Mean height (m)	-0.78%	0.16%	-0.57%	0.97%
Total volume (m^3/ha)	2.16%	2.50%	-0.76%	-2.43%
Mean DBH (cm)	-1.00%	0.50%	-0.35%	-0.34%

Overall, results for species-wise volumes were significantly worse compared to the non-species related attributes (Table 31). Whereas CNNs were the most accurate models before, they were by far the worst performing method for each tree species. It is worth noting, however, that CNN was the only method where we specifically forced the species-wise predictions to add up to the predicted total volume, and there are slight errors with the other models. For instance, the RMSE-% between predicted total volume and sum of species-wise volumes is around 17%. As most of our field plots were pine-dominated, results for pine were the most accurate, while results for the other two species groups were significantly worse.

Table 31: Results for species-wise volumes

	KNN-GA	RF	ANN	CNN
RMSE				
Pine (m^3/ha)	37.08	42.60	41.61	43.87
Spruce (m^3/ha)	23.67	24.61	22.26	51.18
Deciduous (m^3/ha)	18.18	20.65	19.07	26.28
RMSE-%				
Pine (m^3/ha)	39.20%	45.03%	43.99%	46.38%
Spruce (m^3/ha)	114.08%	122.20%	124.11%	140.86%
Deciduous (m^3/ha)	86.25%	97.93%	90.44%	124.64%
BIAS				
Pine (m^3/ha)	1.49	5.11	-5.84	5.15
Spruce (m^3/ha)	0.05	-1.16	0.44	2.85
Deciduous (m^3/ha)	0.25	-2.68	1.42	-4.29
BIAS-%				
Pine (m^3/ha)	-1.60%	5.71%	-6.17%	5.76%
Spruce (m^3/ha)	0.14%	-3.09%	1.20%	8.50%
Deciduous (m^3/ha)	1.21%	-11.30%	6.74%	-16.90%

Overall, utilizing CNNs for predicting these forest variables showed promising results, as they were able to achieve better results for half of the observed stand attributes, even with only low-density LiDAR data. However, we were not able to sufficiently combine aerial imagery with LiDAR point clouds, which resulted poor performance in species-wise volume predictions. Nevertheless, state-of-the-art deep learning methods showed improvement for acquiring structural indicators (e.g. total volume, mean height), whereas other tested methods enable us to predict tree species group distributions in some degree.

Currently, the National Land Survey of Finland (NLS) is updating the nation-wide LiDAR data with higher resolution data (5p/m^2), and it is expected that these data are sufficient for predicting also the horizontal and vertical structure of the forest canopy, one of the key ecosystem indicators, for which the current point density is not sufficient enough. Also, these data should improve also results for species-wise attributes.

3.5 Conclusion and recommendations

This research provided an overview of proposed ecosystem condition indicators for forest ecosystems. We also demonstrated how structural state and landscape indicators can be predicted and calculated with novel remote sensing and machine learning methods, as well as more ‘conventional’ spatial analysis of existing national spatial datasets.

As a rule of thumb, a set of around six to ten indicators for a given ecosystem type generally should provide sufficient information to assess the overall condition of an ecosystem asset, provided they are well selected (UNSD 2020). Ideally, one to two indicators from every SEEA ECT class should be enough, but very challenging to achieve considering the heterogeneity of ecosystem types. Especially the abiotic physical and chemical state indicators are difficult because of lack of spatial data.

The final selection of indicators for Finland is out of scope of this study. The selection should consider not only forests but also other ecosystem types (terrestrial and marine), it should be based on existing ecological knowledge and monitoring systems, with ecologists, remote sensing experts and statisticians directly involved in the selection process.

For forest ecosystems, a consensus should be reached what are the forest ecosystem types that should be included in the Finnish ecosystem type classification and ecosystem accounting. In this study, we took a rather straightforward division by dominant tree species group. The ecosystem services provided by coniferous- and deciduous-dominated forest types are similar but there are also differences which can justify the choices made in this study. Another way of defining forest types would be by tree species composition, since it plays a significant role in maintaining biodiversity: mixed-species forests can host greater species richness and provide more ecosystem services compared to monocultures of conifers (Brockerhoff et al. 2017). Still more approaches include separation between managed forests and forests in natural state, by fertility class, or by forest age groups or development stages.

We demonstrated the use of a robust and scalable modelling framework for calculating landscape indicators for forest fragmentation using existing spatial data. The framework is based on SEBI and previous research done at the JRC and EEA, but for the first time applied in the draft SEEA-EA framework in Finland. By using existing datasets, costly acquisition and processing of new data can be avoided, and time series of MS-NFI and LULC data is available for accounting purposes since 2009. Scalability comes from the fact that the landscape indicators can be applied for any EAA from municipal to country level.

The landscape indicators were implemented with free open source software using models and scripts whenever possible, which enhances repeatability and transparency of the methodology. However, user interaction is still required as the software are based on graphical user interface. The modeling framework could be further automized with R or Python code, which could be the next step forward. More further work exist in definition of reference level values for the indicators (upper and lower levels), and development of connectivity and other indicators (not only landscape) from these datasets.

In this work, we demonstrated that state-of-the-art deep learning methods are able to achieve similar or better results than the method currently in operational use, even with most likely unoptimal voxelization process. As Ayrey et al (2018) had data with considerably higher point density (9-15 pts/m² compared to less than 2 pts/m² used here), the dimensions of the voxel grid are most likely not optimal for our data resulting in mostly empty voxel grids. Nevertheless, CNNs were able to find such features from these data that they were able to outperform other methods for non-species related forest attributes. As the amount of nation-wide LiDAR data increases, the potential for utilizing modern deep learning methods also in operational use becomes a viable alternative. However, LiDAR data alone is not well-suited for estimating tree-species related variables, so either aerial images or satellite images are needed for those attributes. Especially satellite imagery has large potential, as their temporal resolution and spatial coverage are by far the best among the data sources. As shown in Brandt et al (2020), satellite data and deep learning can be used to map even individual trees, though the performance of this method in dense-canopy boreal forests is still an open question. Nevertheless, deep learning has the potential to revolutionalize ecosystem observation and enable more accurate ecosystem indicators.

3.6 Lexicon

CNN	Convolutional neural networks (machine learning method)
DBH	Diameter at breast height
DL	Deep learning (a subgroup of machine learning)
EAA	Ecosystem accounting area
ECT	Ecosystem condition typology
EO	Earth observation
GBM	Gradient boosting trees (machine learning method)
GBR	Gaussian process regression (machine learning method)
KNN-GA	k-nearest neighbors with genetic algorithm (machine learning method)
LAI	Leaf area index
LUKE	National Resources Institute Finland (*Luonnonvarakeskus*)
LULC	Land use / land cover
ML	Machine learning
MS-NFI	Multi-Source National Forest Inventory
MSPA	Morphological Spatial Pattern Analysis
RF	Random forest (machine learning method)
SEBI	Streamlining European Biodiversity Indicators
SEEA-EA	System of Environmental Economic Accounting – Ecosystem Accounting (*draft version*)
SEEA-EEA	System of Environmental Economic Accounting – Experimental Ecosystem Accounting
SVM	Support vector machine (machine learning method)
SYKE	Finnish Environment Institute (*Suomen ympäristökeskus*)

3.7 References

Akar, Ö., Güngör, O. 2012. Classification of multispectral images using Random Forest algorithm. Journal of Geodesy and Geoinformation. Vol.1, Iss.2, pp.105-112

Annala, Leevi; Honkavaara, Eija; Tuominen, Sakari; Pölönen, Ilkka. 2020. Chlorophyll Concentration Retrieval by Training Convolutional Neural Network for Stochastic Model of Leaf Optical Properties (SLOP) Inversion. Remote Sensing 12 2: 22 p.

Ayrey, E.; Hayes, D.J. The Use of Three-Dimensional Convolutional Neural Networks to Interpret LiDAR for Forest Inventory. Remote Sens. 2018, 10, 649. https://doi.org/10.3390/rs10040649

Ballanti, L., Blesius, L., Hines, E., Kruse, B. (2016) "Tree Species Classification Using Hyperspectral Imagery: A Comparison of Two Classifiers" Remote Sens. 8, no. 6: 445. https://doi.org/10.3390/rs8060445

Bishop C. 1995. Neural networks for pattern recognition. New York: Oxford University Press.

Breiman, L.E.O. Random forests. Mach. Learn. 2001, 45, 5–32.

Broadhurst D, Goodacre R, Jones A, Rowland JJ, Kell DB. 1997. Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. Anal Chim Acta. 348(1):71–86.

Brandt, M.; Tucker, C.J.; Kariryaa, A.; Rasmussen, K.; Abel. C.; Small, J.; Chave, J.; Rasmussen, L.V.; Hiernaux, P.; Diouf, A.A.; Kergoat, L.; Mertz, O.; Igel, C.; Gieseke, F.; Schöning, J.; Li, S.; Melocik, K.; Meyer, J.; Sinno, S.; Romero, E.; Glennie, E.; Montagu, A.; Dendocker, M.; Fensholt, R. (2020). An unexpectedyl large count of trees in the West African Sahara and Sahel. Nature. 587, 78-82 (2020). https://doi.org/10.1038/s41586-020-2824-5

Brockerhoff, E.G.; Barbaro, L.; Castagneyrol, B.; Forrester, D.I.; Gardiner, B.; González-Olabarria, J.R.; Lyver, P.O.; Meurisse, N.; Oxbrough, A.; Taki, H.; et al. (2017). Forest biodiversity, ecosystem functioning and the provision of ecosystem services. Biodiversity Conservation 26, 3005–3035.

Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. 2007. Random forests for classification in ecology. Ecology, 88, 2783

Czúcz, B., Condé, S. (2018a). Ecosystem type fact sheet on forests. ETC/BD report to the EEA. https://www.eionet.europa.eu/etcs/etc-bd/products/etc-bd-reports/etc-bd-report-12-2018-fact-sheets-on-ecosystem-condition

Czúcz, B.,Arany, I., Potschin-Young M., Bereczki, K., Kertész, M., Kiss, M., Aszalós, R., Haines-Young, R. (2018b). Where concepts meet the real world: A systematic review of ecosystem service indicators and their classification using CICES. Ecosystem Services 29 (2018) 145–157. https://doi.org/10.1016/j.ecoser.2017.11.018

Czúcz, B., Götzl, M., Schwaiger, E., Sonderegger, G., Condé, S. (2018c). Fact sheets on ecosystem condition: a synthesis. ETC/BD report to the EEA. https://www.eionet.europa.eu/etcs/etc-bd/products/etc-bd-reports/etc-bd-report-12-2018-fact-sheets-on-ecosystem-condition

Estreguil A C., Caudullo, G., Whitmore, C. (2011). Habitat landscape pattern and connectivity indices; Used at varying spatial scales for harmonized reporting in the EBONE project. Wageningen, Alterra, Alterra Report 2297. 80 pp.

Estreguil, C., de Rigo, D., Caudullo, G. (2014). A proposal for an integrated modelling framework to characterise habitat pattern. Environmental Modelling & Software, Vol. 52 (February 2014), pp. 176-191, ISSN 1364-8152. doi: 10.1016/j.envsoft.2013.10.011

European Environment Agency (EEA) (2012). Streamlining European biodiversity indicators 2020: Building a future on lessons learnt from the SEBI 2010 process. EEA Technical report 11/2012. https://www.eea.europa.eu/publications/streamlining-european-biodiversity-indicators-2020/ Accessed on 15th December 2020.

European Environment Agency (EEA) (2015). Fragmentation of natural and semi-natural areas. Indicator Assessment. https://www.eea.europa.eu/data-and-maps/indicators/fragmentation-of-natural-and-semi-1/assessment-1 Accessed on 15th December 2020.

Fix, Evelyn., Hodges, J.L. (1951). Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties (Report). USAF School of Aviation Medicine, Randolph Field, Texas

Franklin, J.F., Forman, R.T.T. (1987). Creating landscape patterns by forest cutting: ecological consequences and principles. Landscape. Ecology 1 (1), 5e18. http://dx.doi.org/10.1007/BF02275261.

Franco-Lopez, H., Ek, A.R. & Bauer, M.E. 2001. Estimation and mapping of forest density, volume and cover type using the k-nearest neighbors method. Remote Sensing of Environment 77: 251–274.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley Publishing Company, Reading, Massachusetts, 412 p.

Gómez, C.; White, J.C.; Wulder, M.A. (2016). Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing. 116, 55-72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Grabska, E.; Frantz. D.; Ostapowicz. K. (2020). Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians. Remote sensing of Environment. 251. https://doi.org/10.1016/j.rse.2020.112103

Haines-Young, R., Potschin, M. (2010). The links between biodiversity, ecosystem services and human well-being. In: In: Raffaelli, D.G., Frid, C.L.J. (Eds.), Ecosystem Ecology: A New Synthesis, vol. 9. Cambridge University Press, Cambridge, pp. 110–139.

Halme, E.; Pellikka, P.; Mõttus, M. (2019). Utility of hyperspectral compared to multispectral remote sensing data in estimating forest biomass and structure variables in Finnish boreal forest. International Journal of Applied Earth Observation and Geoinformation. 83. https://doi.org/10.1016/j.jag.2019.101942

Hastie, T., Tibshirani, R., Friedman, J. (2008). The Elements of Statistical Learning : Data Mining, Inference, and Prediction. New York: Springer. p. 134.

Hinneburg, A., C.C. Aggarwal, and D.A. Keim, 2000. What is the nearest neighbor in high dimensional spaces?, Proceedings of the 26th Very Large Data Bases (VLDB) Conference, 10–14 September, Cairo, Egypt, pp. 506–515.

Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708). https://arxiv.org/abs/1608.06993

Jain, A., Zongker, D., 1997. Feature selection: Evaluation, application, and small sample performance. IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (2), 153–157.

Järnstedt, J., Pekkarinen, A., Tuominen, S., Ginzler, C., Holopainen, M. & Viitala, R. 2012. Forest variable estimation using a high-resolution digital surface model. ISPRS Journal of Photogrammetry and Remote Sensing 74: 78-84.

Kilkki, P., and R. Päivinen, 1987. Reference sample plots to combine field measurements and satellite data in forest inventory, Department of Forest Mensuration and Management, University of Helsinki, Research Notes, 19:210–215.

Kudo, M., and J. Sklansky, 2000. Comparison of algorithms that select features for pattern classifiers, Pattern Recognition, 33:25–41.

Kumar, A., Bi, L., Kim, J., Dagan Feng, D. (2020). Machine learning in medical imaging. Editor(s): David Dagan Feng, In Biomedical Engineering, Biomedical Information Technology (Second Edition), Academic Press, 2020.

Lai, T.-Y., Salminen, J., Jäppinen, J.-P., Koljonen, S., Mononen, L., Nieminen, E., Vihervaara, P., Oinonen S. (2018). Bridging the gap between ecosystem service indicators and ecosystem accounting in Finland. Ecological Modelling 377:51-65. https://doi.org/10.1016/j.ecolmodel.2018.03.006

Lillesand T.M., Kiefer R.W., Chipman J.W. (2004). Remote Sensing and Image Interpretation (5th ed.). John Wiley and Sons Inc, New York, NY, USA.

Maes, J., Teller, A., Erhard, M., Liquete, C., Braat, L., Berry, P., … Bidoglio, G. (2013). Mapping and Assessment of Ecosystems and their Services. An analytical framework for ecosystem assessments under action 5 of the EU biodiversity strategy to 2020. Publications office of the European Union, Luxembourg. http://ec.europa.eu/environment/nature/knowledge/ ecosystem_assessment/pdf/MAESWorkingPaper2013.pdf

Maes, J., Teller, A., Erhard, M., Murphy, P., Paracchini, M.L., Barredo, J.I., … Lavalle, C. (2014). Mapping and Assessment of Ecosystems and their Services: Indicators for ecosystem assessments under Action 5 of the EU Biodiversity Strategy to 2020. 2nd Report. Publications office of the European Union, Luxembourg. http://catalogue.biodiversity.europa.eu/uploads/ document/file/1230/2ndMAESWorkingPaper.pdf

Maes J, Teller A, Erhard M, Grizzetti B, Barredo JI, Paracchini ML, Condé S, Somma F, Orgiazzi A, Jones A, Zulian A, Vallecilo S, Petersen JE, Marquardt D, Kovacevic V, Abdul Malak D, Marin AI, Czúcz B, Mauri A, Loffler P, Bastrup-Birk A, Biala K, Christiansen T, Werner B (2018) Mapping and Assessment of Ecosystems and their Services: An analytical framework for ecosystem condition. Publications office of the European Union, Luxembourg.

Maxwell, A.E.; Sharma, M.; Kite, J.S.; Donaldson, K.A.; Thompson, J.A.; Bell, M.L.; Maynard, S.M. Slope Failure Prediction Using Random Forest Machine Learning and LiDAR in an Eroded Folded Mountain Belt. Remote Sens. 2020, 12, 486.

Modzelewska, A.; Fassnacht, F.E.; Stereńczak. K. (2020). Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data. International Journal of Applied Earth Observation and Geoinformation. 84. https://doi.org/10.1016/j.jag.2019.101960

Mononen L., Auvinen A.-P., Ahokumpu, A.-L., Rönkä, M., Aarras, N., Tolvanen, H., Kamppinen, M., Viirret, E., Kumupla, T, Vihervaara P. (2016). National ecosystem service indicators: Measures of social–ecological sustainability. Ecological Indicators 61: 27–37. http://dx.doi.org/10.1016/j.ecolind.2015.03.041

Moser P, Vibrans AC, McRoberts RE, Næsset E, Gobakken T, Chirici G, Mura M, Marchetti M. 2017. Methods for variable selection in LiDAR-assisted forest inventories. Forestry Int J Forest Res. 90(1):112–124.

Muinonen E., Tokola T. (1990). An application of remote sensing for communal forest inventory. In: Proceedings from SNS/IUFRO Workshop: The Usability of Remote Sensing for Forest Inventory and Planning (conference held in Umeå, Sweden, on 26–28 February). Remote Sensing Laboratory, Swedish University of Agricultural Sciences, Report 4: 35–42.

Mäkisara, K., Katila, M., Peräsaari, J. (2019). The Multi-Source National Forest Inventory of Finland - methods and results 2015. Natural resources and bioeconomy studies 8/2019. Natural Resources Institute Finland, Helsinki, 57 pp. http://urn.fi/URN:ISBN:978-952-326-712-1

Nelder, J.A., and R. Mead, 1965. A simplex method for function minimization, Computer Journal, 7:391–98.

Nemni, E.; Bullock, J.; Belabbes, S.; Bromley, L. Fully Convolutional Neural Network for Rapid Flood Segmentation in Synthetic Aperture Radar Imagery. Remote Sens. 2020, 12, 2532. https://doi.org/10.3390/rs12162532

Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne lidar. Doctoral thesis. Acta Universitatis Agriculturae Sueciae 17.

Pinto, M.M.; Libonati, R.; Trigo, R.M.; Trigo, I.F.; DaCamara, C.C. (2020). A deep learning approach for mapping and dating burned areas using temporal sequences of satellite images. ISPRS Journal of Photogrammetry and Remote Sensing. 160, 260-274. https://doi.org/10.1016/j.isprsjprs.2019.12.014

Pohjankukka, J.; Tuominen, S.; Pitkänen, J.; Pahikkala, T.; Heikkonen, J. 2018. Comparison of estimators and feature selection procedures in forest inventory based on airborne laser scanning and digital aerial imagery. Scandinavian Journal of Forest Research

Pudil, P., Novovicová, J., Kittler, J., 1994. Floating search methods in feature selection. Pattern Recognition Letters 15 (11), 1119–1125.

Pohjola, J., Pekkonen, M., Kopperoinen, L., Kullberg, P., Hurskainen, P, Oinonen, S., Kostamo, K., Jernberg, S., Viitasalo, M., Virtanen, E., Korpinen, S., Kuosa, H., Saikkonen, L., Lankia, T., Lintunen, J., Neuvonen, M., Store, R., Pouta, E., Lai, T., Vihervaara, P. (2019). From existing environmental and economic data and models to experimental accounting of marine, freshwater and forest ecosystem services in Finland (2018-FI-ECOSYSTEMS). Eurostat Grants 2018 final report (unpublished).

Pyo, J.; Duan, H.; Baek, S.; Kim, M.S.; Jeon, T.; Kwon, Y.S.; Lee, H.; Cho, K.H (2019). A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery. Remote Sensing of Environment. 233. https://doi.org/10.1016/j.rse.2019.111350

Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B., Smith, E.R. (2000). Global-scale patterns of forest fragmentation. Ecological Society 4 (2), 3. http://www.consecol.org/vol4/iss2/art3/

Riitters, K., Wickham, J., Wade, T. (2009). Evaluating anthropogenic risk of grassland and forest habitat degradation using land-cover data. Landscape Online 13, 1e 14. http://dx.doi.org/10.3097/LO.200913

Sabat-Tomala, A.; Raczko, E.; Zagajewski, B. Comparison of Support Vector Machine and Random Forest Algorithms for Invasive and Expansive Species Classification Using Airborne Hyperspectral Data. Remote Sens. 2020, 12, 516.

Siedlecki, W., Sklansky, J., 1989. A note on genetic algorithms for large scale feature selection. Pattern Recognition Letters 10 (5), 335–347.

Soille, P., Vogt, P. (2009). Morphological segmentation of binary patterns. Pattern Recognition Letters 30 (4), 456e459. http://dx.doi.org/10.1016/j.patrec.2008.10.015

Stephens, T. Titanic: Getting Started With R—Part 5: Random Forests. Available online: http://trevorstephens.com/kaggle-titanic-tutorial/r-part-5-random-forests/ (accessed on 14 March 2017).

Tokola T., Pitkänen J., Partinen S., Muinonen E. (1996). Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. International Journal of Remote Sensing 17(12): 2333–2351. http://dx.doi.org/10.1080/01431169608948776.

Tomppo, E., 1991. Satellite image-based national forest inventory of Finland. International Archives of Photogrammetry and Remote Sensing 1991 (28), 419–424.

Tomppo, E., Haakana, M., Katila, M., Peräsaari, J. (2008). Multi-Source National Forest Inventory Methods and Applications. Springer. Series: Managing Forest Ecosystems, Vol. 18 2008, XIV. 374 p.

Tomppo, E., Katila, M., Mäkisara, K., Peräsaari, J. (2013). The Multi-source National Forest Inventory of Finland – methods and results 2009. Working Papers of the Finnish Forest Research Institute 273. 216 s. ISBN 978-951-40-2428-3 (PDF). Available at: http://www.metla.fi/julkaisut/workingpapers/2013/mwp273.htm

Tuominen S., Haapanen R. (2013). Estimation of forest biomass by means of genetic algorithm-based optimization of airborne laser scanning and digital aerial photograph features. Silva Fennica vol. 47 no. 1 article id 902. Available at: https://doi.org/10.14214/sf.902

Tuominen, S.; Pitkänen, T.; Balazs, A.; Kangas, A. (2017). Imporoving Finnish Multi-Source National Forest Inventroy by 3D aerial imaging. Silva Fennica 51 (4). https://doi.org/10.14214/sf.7743

Tuominen, Sakari; Näsi, Roope; Honkavaara, Eija; Balazs, Andras; Hakala, Teemu; Viljanen, Niko; Pölönen, Ilkka; Saari, Heikki; Ojanen, Harri. 2018. Assessment of Classifiers and Remote Sensing Features of Hyperspectral Imagery and Stereo-Photogrammetric Point Clouds for Recognition of Tree Species in a Forest Area of High Species Diversity. Remote Sensing 10 5: 28.

Tuominen S., Balazs A., Kangas A. (2020). Comparison of photogrammetric canopy models from archived and made-to-order aerial imagery in forest inventory. Silva Fennica vol. 54 no. 5 article id 10291. Available at: https://doi.org/10.14214/sf.10291

United Nations (2021). System of Environmental Economic Accounting. Frequently asked questions. Available at: https://seea.un.org/content/frequently-asked-questions. Accessed 13.1.2021.

United Nations Statistics Division (UNSD) (2020). System of Environmental-Economic Accounting—Ecosystem Accounting. Draft for the Global Consultation on the complete document. Available at: https://seea.un.org/content/global-consultation-complete-draft

Vihervaara, P., Järvenpää, E., Mattsson, T., Mononen, L., Kortelainen, P., Oinonen, S., Saikkonen, L., Holmberg, M., Akujärvi, A., Bergström, I., Latukka, A., Regina. K., Vanhala, P., Forsius, M., Lintunen, J., Pohjola, J., Laturi, J., Neuvonen, M., Lankia, T., Sievänen, T., Torvelainen, J., Pouta, E. (2018). Towards ecosystem accounting based on innovations and insights on natural capital knowledge. Eurostat Grants 2017 final report. Available at: https://circabc.europa.eu/sd/a/2403d31e-e57f-44f3-85b6-892d89986d7a/FI-2017-Ecosystem%20accounts.pdf

Viinikka, A.; Hurskainen, P.; Keski-Saari, S.; Kivinen, S.; Tanhuanpää, T.; Mäyrä, J.; Poikolainen, L.; Vihervaara, P.; Kumpula, T. Detecting European Aspen (Populus tremula L.) in Boreal Forests Using Airborne Hyperspectral and Airborne Laser Scanning Data. Remote Sens. 2020, 12, 2610. https://doi.org/10.3390/rs12162610

Wall, M., 1996. GAlib: A CLibrary of Genetic Algorithm Components, Version 2.4 Documentation, Revision B, Massachusetts Institute of Technology, 101 p.

Wang et al., 2019. "3D Point Cloud Analysis and Classification in Large-Scale Scene Based on Deep Learning," in IEEE Access, vol. 7, pp. 55649-55658, 2019, doi: 10.1109/ACCESS.2019.2909742.

Whitney, A.W., 1971. A direct method of nonparametric measurement selection. IEEE Transactions on Computers 20 (9), 1100–1103.

Wickham, J.D., Nortonm D.J. (1994). Wickham, J.D., Norton, D.J., 1994. Mapping and analyzing landscape patterns. Landscape Ecology 9 (1), 7e23. http://dx.doi.org/10.1007/BF00135075

Willighagen E., Ballings M. (2015). Genalg: R based genetic algorithm. https://cran.r-project.org/web/packages/genalg/index.html

Öhman, K., Eriksson, L.O. (1998). The core area concept in forming contiguous areas for long-term forest planning. Canadian Journal of Forestry Research 28: 1032-1039.

4 Valuation of wooded land (WP3)

Jani Laturi[1], Johanna Pohjola[2]

[1] National Resources Institute Finland (Luke)

[2] Finnish Environment Institute (SYKE)

In this section we focus on the valuation of wooded land. We provide ecosystem monetary asset accounts for forest capital by using different approaches, namely stumpage value, property price value and net present value methods. Some of the approaches include only the value of timber resources while some include also the bareland value. SEEA Central Framework provides guidelines for valuing forest as an asset (UN et al.2014). Central Framework covers forest accounts both in physical and monetary terms, for which we focus on the latter one.

4.1 Scope and definitions for valuation of forests with monetary asset account

Monetary asset account for timber resources consists in measuring the value of the opening and closing stock of timber resources and the changes in the value of the stock over an accounting period. Monetary asset account records the additions and reductions that appear also in physical account. SEEA guidelines find it important to distinguish between the changes due to the changes in volumes and changes in prices. To provide both effects separately, monetary asset account includes the entry called revaluation that records the impact of price changes during an accounting period.

Monetary asset account covers only those timber resources that have economic value, according to SEEA Central Framework (UN et al.2014). On the other hand, physical accounts cover all timber resources as they may provide other benefits. Timber resources outside monetary accounts include the forests that are in areas in which logging operations are restricted or prohibited, that are in areas that are inaccessible or remote and hence logging is not economically viable or do not belong to a commercially useful species. In addition, SEEA guidelines emphasize the distinction between cultivated and natural forests. The growth in cultivated timber resources is considered as a process under the direct control, responsibility and management of institutional units. Consequently, the growth is recorded as occurring within the production boundary. The growth of natural timber resources, on the other hand, is not considered to take place within the production boundary and is recorded as entering the production boundary only at the time the tree is removed from the forest or other land area. In the case of Finland, most of the timber resources are commercially utilized and under forest management.

Resource rent on timber resources can be derived as the gross operating surplus from the harvest of timber resources (after taking into account specific taxes and subsidies) less the value of the user costs of produced assets used in the harvesting process. Resource rent can be estimated by using estimates of the stumpage price, which is the amount paid per cubic meter of timber by the harvester to the owner of the timber resources. The harvesting costs should include felling costs as well as costs of thinning, other management costs and rent on land.

4.2 The value of forests and timber resources in Finland by using three different approaches

In here we compare three different method to produce monetary asset account for forestry land in Finland for period 2010-2014.

a) Stumpage value approach

b) Property price approach

c) Net present value approach

4.2.1 Stumpage value approach

In Finland forest and forestry statistics are maintained by Natural Resources Institute Finland (Luke). The statistics present investment returns on wood production in non-industrial private forests (NPF), which includes calculation for the asset value. The forest asset value is called stumpage value. It is calculated by multiplying the assortment-specific volumes of forests by their standing sales prices (Luke 2021). This corresponds the situation in which forests are harvested instantaneously. Stumpage value approach do not include the value of bareland or expectation values for standing trees.

Table 32: Monetary asset account for timber resources in forestry land available for wood production with stumpage value approach

Period	2010-2014
(billion €)	
Opening stock of timber resources	60.75
Natural growth	14.07
Removals	9.12
Revaluation	-0.65
Closing stock of timber resources	65.05

In the Table 32 the monetary asset account with values of opening and closing stock is presented based on the stumpage value approach. The accounting area applied is forest land available for wood production as in other approaches. We have used detailed data provided by Luke on the regional stumpage prices and volumes by wood assortments to calculate the value of timber resources. This data that is used in Luke to calculate forestry investment returns as described above, covers only the non-industrial private forests available for wood production. Thus, we have to convert the obtained forest value for all forest land available for wood production[9]. Natural growth describes the value of growth and transition of assortment-specific volumes. Removals are the value of stumpage earnings during the period. The revaluation effect describes the change in the price of the resource multiplied by the quantity at the end of the period.

4.2.2 Property price approach

The real estate purchase price register is maintained by the National Land Survey of Finland. In taxation in Finland the purchase price data of forest properties is used for valuation of forest properties which value has not been specified otherwise. Taxation administration publishes each year a table of these

[9] The value for timber resources for non-industrial private forests available for wood production has been converted to value for all forest land available for wood production by multiplying it by 1.43 that is the relation between the areas of all forest land and private forest land available for wood production in 2004-2008, which is the most recent period this information has been reported in Finnish forest statistics.

values, which are based on statistics for the previous three years (Verohallitus 2010, Verohallitus 2015). These purchase prices include both the value of the land and the value of the stand. In addition, other forest ecosystem services affect the price of the forest.

Table 33: Monetary asset account for forests and timber resources in forestry land available for wood production with property price approach

Period	2010-2014
(billion €)	
Opening stock of timber resources	37.23
Natural growth	9.12
Removals	9.12
Revaluation	9.99
Closing stock of timber resources	47.22

In the Table 33 the values of opening and closing stock based on the property price value approach are presented. The value of volume growth of forest is equal to removals because the revaluation includes both the stumpage and bareland value changes of forests. Removals are the value of stumpage earnings during the period. The revaluation effect describes the change in the price of the forest land multiplied by the quantity (hectares) at the end of the period. In here, the forestry land available for wood production is assumed constant and equal the land area in the net present value approach and in the FinFEP-model.

4.2.3 Net present value approach

In the forest economy the value of forest land in timber production is calculated with so called Faustmann formula. Faustmann formula does not take into account other ecosystem services than timber production. Typically, other ecosystem services such as berry picking, mushroom picking, and recreational use benefit from forest aging. Thus, taking into account forest ecosystem services lengthens forest rotations. In the case of Finland, this is reflected in the fact that the preferences of forest owners must be taken into account in order to explain the harvest volumes and the existence of old-growth forests (Lintunen et al. 2015). In the SEEA Central Framework this challenge has been identified: “The primary difficulty in applying these NPV approaches lies in the extent to which information is available on the age structure of the trees and how these trees will mature into the future.”. However, methods such as remote sensing are available to produce the age structure, but methods for determining mature age and its relationship to forest owners’ preferences requires development.

In here, we produce monetary asset account for forest timber production based on the estimated future harvests. The future harvest volumes are results of Finnish forest and energy sector model (FinFEP; Lintunen et al 2015). FinFEP model captures the development of age classes in Finnish forests. We use the version of FinFEP-model which is calibrated to year 2010 and run simulation for 165 years. Discount rate of 2.7 % is used that is consistent with the one used in harvesting decisions in the model. Because the simulated time is finite the estimated value is underestimation also in the case of wood production. The preferences of forest owners are assumed to be constant. So, the harvest rules are unchanged. However, estimates of demand in the forest and energy sectors affect production and investment in these sectors. Demand for end products in the forest sector is assumed to grow by 1% annually, with the exception of paper products that is assumed to decrease 1% per year. In addition, given the aging of forest resources, estimating future harvest volumes requires extensive modeling. Timber prices are determined in the model. They vary over accounting period, balancing the supply and demand of timber. Thus, our approach is market-based.

Table 34: Monetary asset account for forests and timber resources in forestry land available for wood production with NPV approach

Period	2010-2014
(billion €)	
Opening stock of timber resources	90.29
Natural growth	9.16
Removals	9.97
Revaluation	
Closing stock of timber resources	89.48

Here, forest owners' preferences and other ecosystem services affect wood production and its discounted value. However, we are unable to estimate the value of those other ecosystem services at this time. In the Table 34 the value of opening and closing stocks are discounted net profits of harvests during the simulation. Removals are discounted net profits of harvest during the period 2010-2014. Because opening and closing stocks are based on same simulation the natural growth balance the values of the accounts i.e. natural growth equals removals plus the difference between closing and opening stock.

4.3 Value of carbon related forest ecosystem services

Carbon related ecosystem services are the most straightforward to quantify in addition to timber production as their volume is directly related to timber and there exist suitable market prices for valuation. SEEA revision paper on defining and valuing the carbon related ecosystem services (Edens et al. 2019) recommends that carbon sequestration should be considered as the (only) carbon related ecosystem service. On the other hand, the amount stored (as a result of sequestration or pre-existing) is a stock variable that may be included in the condition account and/or carbon account.

4.3.1 Definitions for carbon sequestration

Edens et al (2019) argue that the definition and measurement of carbon sequestration should be aligned with / be complementary to IPCC guidelines as much as possible. Thus, they recommend that carbon sequestration should be measured as net ecosystem carbon balance (NECB) that accounts also carbon losses. This approach is also suggested in Lintunen et al (2018).

Net ecosystem carbon balance is defined as

NECB (net ecosystem carbon balance) = Net Biome production (NBP) = NEP – Carbon loss from Disturbance/Land-clearing/Harvest

in which

NEP (net ecosystem production) = NPP – soil respiration = GPP – ecosystem respiration

There are two approaches to analyze the supply of carbon sequestration by ecosystems. In stock-difference method sequestration is derived by comparing changes in stocks of carbon over time, for instance on the basis of forest inventories and soil carbon measurements. This approach may be called an indirect method, as sequestration can be derived as a residual. Gains-Loss method estimates carbon sequestration directly (called in IPPC guidelines) and involves the quantification of all key inflows and outflows of carbon per ecosystem unit.

4.3.2 Valuation of carbon sequestration

Edens et al (2019) identify four options to value carbon, namely social cost of carbon (SCC), abatement cost approach, observed market prices and consumer preferences, and discuss the benefits and deficiencies of these approaches. They recommend the use of ETS prices where they are available, and in countries with an ETS they recommend using these ETS prices also as “best available estimates” for those sectors which are not covered by the respective ETS. A major advantage of ETS market prices is that they provide a market-based exchange value that is aligned with the SNA (and SEEA) exchange value notion. In addition, they are readily available. Beyond this, many transactions concerning carbon are already recorded in the national accounts. However, they suggest that ETS based carbon valuations should be complemented with an additional valuation based at a (global) SCC estimate.

4.3.3 The net present value of carbon sequestration in Finland

Value of carbon sequestration in Finland is presented as net present value with calculation period of 165 years. The amounts of carbon sequestration are based on the scenario we utilize for NPV calculation for timber production described above. The carbon sequestration is calculated as changes in the carbon stock on the above and below ground biomass of trees. According to Edens et al (2019), we value carbon by using the price of European Union Allowances (EUA) in the European Union Emissions Trading Scheme (EU ETS) and estimate for social cost of carbon. We have chosen carbon prices of 15, 30, 50 and 100 €/t CO_2 to illustrate the sensitivity of the value of carbon sequestration to price of carbon. The price of 15 €/t CO_2 is the average price of EUA for year 2010 that is the first year of the accounting period while the price of 30 €/t CO_2 is close to the current price of EUA. The price of 50 €/t CO_2 is the carbon price in the EU in 2040 according to the reference scenario published by European Parliament (2016). The upper limit for estimates of social cost of carbon is reflected in the carbon price of 100 €/t CO_2 (Lintunen et al 2018). Carbon price is assumed to be constant over time. NPV for carbon sequestration is presented in Table 35 for opening stock in 2010 and closing stock in 2014.

Table 35: Net present value of carbon sequestration with different prices of carbon, billion euros

Price of EUA, €/t CO_2	Value of opening stock	Value of closing stock
15	18.1	17.5
30	35.8	35.0
50	59.6	58.4
100	119.2	116.8

Net present value for carbon sequestration varies between 18–119 billion euros depending on the price of carbon used. The amount of carbon sequestration is largest in the beginning of the NPV calculation period of 165 years and is considerably reduced over time.

4.4 Conclusions

We have provided the value of forests for Finland by using three different approaches, namely stumpage value, property price value and NPV approaches. In net present value approach, the forest-energy sector model FinFEP was utilized in order to produce the future harvest levels and capture the development of age classes in forests. The value of forests based on timber production was 37–90 billion euros, depending on the approach used. The higher value in the case of NPV approach is partly explained by the fact that forests are rapidly growing in the future. We also calculated the value of carbon sequestration with NPV approach. The value was estimated to be 18–120 billion euros for carbon prices of 15–100 €/t CO_2 based on carbon price in EU ETS and social cost of carbon. With most likely carbon prices, 30 and 50 €/t CO_2 the value of sequestration was 36 or 60 billion euros that is notable compared to value of timber production.

The forest asset values consist on the value of bareland, net present value of standing trees and net present value of all ecosystem services. As noted above, they cannot be properly valued by one method. According to forest economics, the NPV method is scientifically correct, but its implementation requires the use of a model and the making of predictions. In any case, there are currently no methods for valuing multiple simultaneous ecosystem services. Property prices includes also values related other ecosystem services than wood production. In principle, it offers the possibility to value different ecosystem services, but the amount of data may limit its use.

4.5 References

Edens B., Elsasser P., Ivanov E. (2019). Discussion paper 6: Defining and valuing carbon related services in the SEEA EEA. Paper submitted to the Expert Meeting on Advancing the Measurement of Ecosystem Services for Ecosystem Accounting, New York, 22-24 January 2019 and subsequently revised. Version of 15 March 2019. Available at: https://seea.un.org/events/expert-meeting-advancing-measurement-ecosystem-services-ecosystem-accounting

European Parliament (2016). Draft Report of the Committee on the Environment, Public Health and Food Safety. Compromise Amendments 1-17 on the Proposal for a Directive of the European Parliament and of the Council amending Directive 2003/87/EC to Enhance Cost-effective Emission Reductions and Low-carbon Investments. 2015/0148(COD), 14 December 2016

Lintunen, J., Laturi, J., & Uusivuori, J. 2015. Finnish Forest and Energy Policy model (FinFEP): a model description. Natural resources and bioeconomy studies 59/2015. http://urn.fi/URN:ISBN:978-952-326-119-8

Lintunen, J., Pohjola, J., Laturi, J. (2018). Valuation of forest related ecosystem services. In Vihervaara et al. (2018): Towards ecosystem accounting based on innovations and insights on natural capital knowledge. Eurostat Grants 2017 final report. https://circabc.europa.eu/sd/a/2403d31e-e57f-44f3-85b6-892d89986d7a/FI-2017-Ecosystem%20accounts.pdf

Luke. 2021. Forestry as an investment. Natural Resources Institute Finland (Luke). https://stat.luke.fi/en/quality-description-forestry-investment_en

United Nations, European Union, FAO, IMF, OECD, World Bank (2014). System of Environmental- Economic accounting 2012: Central framework. UN, New York. https://unstats.un.org/unsd/envaccounting/seearev/seea_cf_final_en.pdf

Verohallitus (2010). Varojen arvostaminen perintö- ja lahjaverotuksessa. https://www.vero.fi/contentassets/9ec0fac42cd046f5897f7ecd69faf85f/varojen_arvostaminen_perint%C3%B6-_ja_lahjaverotuksessa_p%C3%A4ivitys_1_9_2010.pdf

Verohallitus (2015). Varojen arvostaminen perintö- ja lahjaverotuksessa. https://www.vero.fi/contentassets/47eb2a04d85042c6926000d4d1c79957/varojen-arvostaminen-perinto--ja-lahjaverotuksessa-vuosi-2016.pdf