



cross forest



Co-financed by the European Union
Connecting Europe Facility

User Specifications and Requirements

Deliverable D3.1 Public
v2.0

Keywords: Forest fire, Wood quality, modellization, simulation

CROSS Harmonization & HPC modelization of FOREST Datasets

Action number: 2017-EU-IA-0140

Agreement INEA/CEF/ICT/A2017/1566738

Table of Contents

1	Introduction.....	5
1.1	Purpose.....	5
1.2	Scope.....	5
2	Cross-Forest Platform.....	6
2.1	Linked Open Data in Cross-Forest.....	6
2.2	HPC.....	7
3	Pilots specifications and requirements.....	9
3.1	CAMBRIC: CALidad de la Madera en Bosques mlxtos.....	9
3.1.1	Overall Perspective.....	9
3.1.2	User interface.....	11
3.1.3	Input Data.....	11
3.1.4	Data transformation.....	14
3.1.5	Outputs.....	16
3.2	FRAME: Forest fiRes Advanced ModElization.....	18
3.2.1	Overall Perspective.....	18
3.2.2	User interface.....	19
3.2.3	Input data.....	20
3.2.4	Data transformation.....	21
3.2.5	Outputs.....	22
4	Conclusions.....	23
	References.....	24
	Definitions, acronyms and abbreviations.....	26

List of Figures

Figure 1 - CAMBRIC pilot overall perspective.....10

Figure 2 - High scenario for SIMANFOR.....17

Figure 3 - ERVIN simulation screenshots.....19

Figure 4 - EMERCARTO platform screenshot.....20

List of Tables

Table 1 – CAMBRIC pilot overall perspective.....9

Table 2 - Spanish Forest Map (MFE) variables in *Poligon* data-sheet.....12

Table 3 - Study area variables.....12

Table 4 - Stand characteristics variables.....13

Table 5 - Tree level variables from NFI.....13

Table 6: Taxa to be checked.....15

Table 7 – FRAME pilot overall perspective.....18

Table 8 – External datasets to be considered as input data for FRAME pilot.....21

Table 9 – FRAME pilot overview of questions to be answered.....21

Contractual Date of Delivery to the EC:		May 2019	
Actual Date of Delivery to the EC:		19/06/2020 (v2.0)	
Editor(s):		Felipe Bravo (Uva) Cristóbal Ordóñez (Uva) Victor Gonzalvo (TRAGSA)	
Contributor(s):		Álvaro Fanego (SCAYLE) Alexandra Fonseca (DGT) Belén Fierro (TRAGSATEC) Daniel Molina (TRAGSA) Asuncion Roldán (TRAGSA)	
DOCUMENT HISTORY			
Version	Version date	Responsible	Description
0.1	29/05/2019	UVa	Draft for Uva TRAGSA comments
0.2	30/05/2019	UVa	Draft for Cross-Forest consortium comments
0.3	31/05/2019	UVa	Draft for Cross-Forest consortium comments
1.0	03/06/2019	UVa	Version to INEA
1.1	05/05/20	UVa	Draft for requirements improvement
1.2	01/06/20	UVa	Draft for Cross-Forest consortium comments
1.4	15/06/20	UVa	Draft for Cross-Forest consortium evaluation
1.5	17/06/20	UVa	Draft ready for final edition
2.0	30/06/20	UVa	Version to INEA

The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

Copyright © 2020, Cross-Forest Consortium.

1 Introduction

This document defines the procedures developed in the Pilots and practical scenarios of Cross-Forest, FRAME and CAMBRIC, aimed to design and develop the practical use cases for implementation of HPC services applied to forestry environment.

This second version of the deliverable updates the first version of the pilots requirements, enhancing and completing the ones included in deliverables 3.1 v1.0 and 3.2.

For each pilot, information about the available data sources will be provided, the sequence of steps to be carried out, the intended outputs and the necessary transformations of the data sources.

The utilization of cross-border forest datasets in FRAME scenario represents a fundamental progress in cross-border information sharing in Europe, allowing better forest management and protection against fires and other hazards.

The proposed pilots will use datasets gathered by the Directorate General for Territory – DGT (Portugal) and the Ministerio para la Transición Ecológica y el Reto Demográfico – MITERD (Spain) to create harmonized data models oriented towards common publication of forest information.

In this sense, it is essential to define what processes are involved and who should participate. In order to achieve the main objective of the project, we must capture the sequence of actions performed or to be carried out by an external entity on the system to achieve a quantifiable goal.

1.1 Purpose

The main purpose of this document is to define the Cross-Forest platform and describe how pilots interact with the dataset, in the Cross-Forest Project scenarios, to ensure that the endpoint repositories, external data sources and sequences can answer the project team questions regarding forest management and fire spreading scenarios.

Both pilots, FRAME and CAMBRIC, aim to use two previously developed simulators ERVIN and SIMANFOR, but adapted to an HPC environment so as to enhance their capabilities. In both cases, software was built for a Windows platform environment in computer or server deployment, so it is compulsory to adapt their cores to a multiplatform environment able to run in HPC. There were also a prior input and output database structures designed for a PC or server stage that should be transformed into a Linked Open Data dataset, as far as possible.

1.2 Scope

In the first stage of the project the teams responsible of each pilot met SCAYLE team in order to discuss different options to tune or transform the simulators and make them work in Caléndula/HPC environment. Provided both pilot teams have already achieved a suitable simulator, the scope of this document consists of:

- a) To guide the project development, through the identification of information required and how to make the necessary transformations in each pilot.
- b) To validate the project development, through the pilots, identifying the required sources, formats and expected outputs.
- c) To identify sources (internal and external data) that allow obtaining the expected data in each test.
- d) To guarantee the access to internal and external sources and data, that will allow obtaining the expected data.

2 Cross-Forest Platform

Cross-Forest Platform aims to provide necessary data sources for modelling forest fire behaviour and evolution as well as modelling forest growth and yield. The two pilots involved in the project will probe how to perform simulations with these datasets, but the platform is intended to serve lots of other uses in the forestry field, and is envisioned to be fed with other sources of data in the future, such as regional inventories, or other kinds of data such as LiDAR or Satellite.

To achieve this exigency there are two keystones of Cross-Forest platform which will help to configure all our requirements: (i) well designed datasets and its ontologies, and (ii) High Performance Computing, both explained herein below.

2.1 Linked Open Data in Cross-Forest

All data required for each pilot will be detailed in the next chapter, but Cross-Forest platform aims to serve as a Digital Service Infrastructure (DSI) providing Open and Linked Data Sets which have been collected by public bodies, and a harmonised access must be granted. The general description of datasets and ontologies contained in the Cross-Forest DSI will be provided in this section.

Cross-Forest platform will be comprised by three main datasets: i) National Forest Inventory (NFI), ii) Forest Map and iii) Spatial Grid. All of them will be accessible as Linked Open Data for which original data should be transformed into RDF format. All details are described in Deliverable 2.2 of the project.

Both, Spanish and Portuguese NFI would be included in the Platform, and by now the 3rd edition of the Spanish data is already deployed. The second edition is in the process of being deployed and from the fourth edition, still in measurement phase, every province will be included as LOD in the Platform as the data processing is finished. For Spanish NFI, it is planned to keep the last version of the data accessible for every province, and maintain the previous one until the new one is released. The Portuguese Forest Inventory to be stored in Cross-Forest platform is the 6th National Forest Inventory (IFN6). One of its sampling processes corresponds to measurements and evaluations of vegetation in the field, carried out on a set of about 12 thousand sampling points, distributed throughout the territory and performed during the year 2015. Based on the data collected in these sampling points, several biometric parameters of vegetation are estimated.

Both Spanish and Portuguese Land Use Maps will be deployed. For the moment, the published dataset includes the Spanish Forest Map (2nd edition at scale 1:50,000), and the 3rd edition (at scale 1:25,000) is being prepared. Similarly to Spanish NFI, it is planned to keep the last version of the Forest Map accessible, releasing available data and maintain the previous version in the meanwhile. From Portugal the National Land Use Land Cover Map (COS) for 2018 (DGT), completed at the end of November 2019 and available through visualization (WMS) and downloading (ATOM) services will be used. COS is a map of polygons, at the 1:25000 scale, with a defined minimum cartographic unit of 1 ha with a distance between lines greater than or equal to 20 m. The COS nomenclature was reformulated for the production of COS2018 and the nomenclature now includes 83 classes.

Concerning Portuguese data, next steps in the project include creation of new ontologies for the IFN (Inventário Florestal Nacional) and COS (Carta de Uso e Ocupação do Solo de Portugal Continental), as well as an upper ontology able to describe transnational data. Afterwards, IFN and COS data can be transformed into RDF and included in Cross-Forest platform.

The third dataset of Cross-Forest Platform, Spatial Grid, is able to provide a raster representation of positions and geometries over the territory. It is being prepared, and details are described in Deliverable 2.2.

As a necessary germinal step, several ontologies have been developed in the project, detailed in Deliverable 2.1. Apart from one ontology for each of the described datasets, three complementary and necessary ontologies have been developed for:

- Simple Positions Ontology: to represent positions of entities.
- EPSG ontology: to disentangling relationships among different Coordinate Reference System in which different datasets are originally referenced.
- Simple Measures Ontology: to characterize and transform measures of variables as necessary, depending on required units for their final use.

2.2 HPC

High Performance Computing is important for both pilots, (i) in the case of FRAME because predictions about real fire should be achieved as fast as possible to fight fires, and because of the need of performing multiple simulations under different configurations, and (ii) in the case of CAMBRIC because simulations are performed with a big dataset and imply complex calculus which take long time to be solved in PC computers.

Cross-Forest Platform is hosted in Caléndula Supercomputer (located in SCAYLE), for which important resources have been devoted, as explained in Deliverable 1.1, comprising 2 powerful virtual machines in which Virtuoso software has been installed to give querying services for datasets, SSH server for securing connections and enough room storage to host the datasets transformed and produced by the project.

Apart from the Project Platform, SCAYLE is dedicating High Performance Computing resources to the Pilots, also described in Deliverable 1.1. For CAMBRIC, the simulator runs over Python (reprogrammed from original C#), and it was prepared to parallelize over Dask, but it was found embarrassingly parallel, which means that the result of a simulation for each plot is completely independent of the simulation for any other plot, so simulations will

be parallelized by splitting input data into groups of plots without any effect in output results. On the other hand, FRAME simulations run over C/C++ (ported from original C#), so will smoothly perform in Caléndula without any further installation provided C is natively supported in Linux systems.

The improvement in performance for both pilots provided by the use of HPC resources is being checked through test executions in Caléndula. Tests included in the first iteration are explained in Deliverable 3.2 (CAMBRIC in section 2.3.1 and FRAME in section 3.3.1). Some more details about results from the first iteration can be found in Deliverable 4.3 where a comprehensive evaluation plan for the project is fine-grained. As a results from the delay in the adaptation of simulators to HPC, some of the planned tests are currently on-going. The tests will continue in an iterative process until the end of the project, and results will be included in coming progress and technical reports.

3 Pilots specifications and requirements

Although Cross-Forest dataset will be exploited by both pilot projects, some peculiarities of data requirements for each one make it advisable to describe them separately. National Forest Inventories and Forest Maps are the core data to be used by pilots, so available data from each country (Portugal and Spain) will be stored in the project platform. Additionally, both pilots should use other sources of data such as climate, soil, or topographic, which will be adequately explained in the following chapters.

3.1 CAMBRIC: CALidad de la Madera en Bosques mixtos

3.1.1 Overall Perspective

The aim of CAMBRIC is to forecast wood quality in mixed forest on a large scale (over municipal/regional areas). Using Big Data as source information provided by National Forest Inventory, a set of yield models adequate for each stand, and different management scenarios, we will develop a dynamic information platform with wood production for different end uses.

There will be tested the five silvicultural scenarios proposed by Duncker et Al. (2012): (1) Passive: Unmanaged forest nature reserve; (2) Low: Close-to nature forestry, (3) Medium: Combined objective forestry, (4) High: Intensive even-aged forestry and (5) Intensive: Short-rotation forestry, always depending on the suitability of the species composition into natural area under study to each silvicultural management schema.

Table 1 – CAMBRIC pilot overall perspective

DESCRIPTION	Use of an HPC-adapted simulator for predicting stand evolution and wood quality production under different silvicultural regimes, using data from Spain (Forest Map and Forest Inventory).	
PRECONDITION	The simulator is running in an HPC environment. The datasets of the Project platform are accessible and published. The external databases are published/available. Access to the Endpoint repository and external databases is available.	
NORMAL SEQUENCE	STEP	ACTION
	1	Select taxa based on ecological conditions described for Spain.
	2	Select and program all appropriate silvicultural scenarios from Duncan options.
	3	Select National Inventory plots representing the area for selected taxa.

	4	Run all suitable scenarios for the selected taxa.
	5	Publish all scenarios and outputs into project platform.
POSTCONDITION		<p>Main taxa areas selected based on Spanish Forest Map.</p> <p>Different adequate silvicultural regimes set up through management scenarios and new simulations can be obtained.</p> <p>Geographical scope of simulations comprehends the main forested areas of Spain.</p> <p>Publication of results as Linked Open Data.</p>

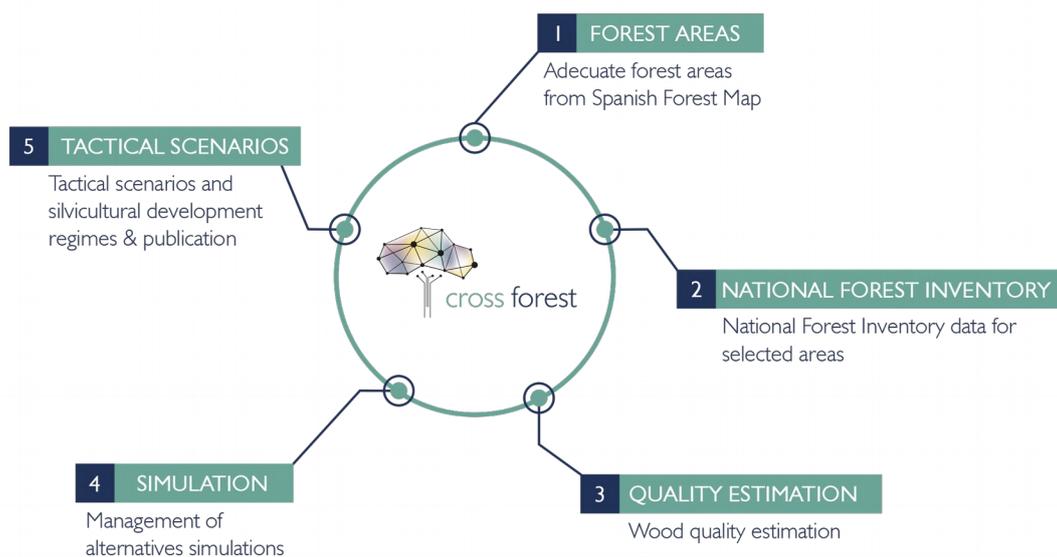


Figure 1 - CAMBRIC pilot overall perspective

These management schemes will be simulated across a gradient of forest mixtures (from pure to multi-species stands) for the target species. Simulations will allow to classify timber products on a log size and stem knot-free length basis. The work-flow (see Figure 2) will be (i) select adequate areas with the information provided by the Spanish Forest Map covers, (ii) extract forest inventory data for the selected areas, (iii) estimate wood quality by applying stem taper equations to inventory data, (iv) simulate different silvicultural alternative for the Spanish NFI data to explore the impact of species mixture on wood quality and (v) develop tactical scenarios for the different silvicultural regimes and publish output data following the Linked and Open data approach defined in the project.

The following taxa will be selected as model species to analyse the availability of wood and its quality for different end uses: *Pinus sylvestris*, *Pinus pinaster*, *Pinus nigra*, *Fagus sylvatica*, *Quercus pyrenaica*, *Quercus robur* and *Quercus petraea*, for pure stands and for the found admixtures stands of these species.

CAMBRIC pilot is based on scenario simulation, so all requirements are related with data needed for area selection and input variables of simulations, and with the equations ruling this simulations (See Deliverable 3.2 Annex I). Regarding output, traditional yield tables and

industry sizes and qualities are the main issues. All these requirements, including user interfaces will be following described.

3.1.2 User interface

For all of the different steps involved in the CAMBRIC pilot, there are different tools which should be engaged, and are next detailed.

Forest area adequate to each taxa selection, as well as NFI plots selection will be obtained based on SPARQL queries, provided both datasets are already included in Cross-Forest platform. The help of Forest Explorer (<https://forestexplorer.gsic.uva.es/explorer/>) to visualize data involved areas will be very useful.

Quality estimation depends on equations and wood size dimensions that will be included in the models of each taxa, and will be performed in scenario simulations with SIMANFOR (www.simanfor.es).

Simulations will be carried out with a new version of SIMANFOR (www.simanfor.es) in HPC environment. Each of the scenarios can provide a new dataset for each area, suitable to be included in the Cross-Forest platform, and further analysis and visualization with Forest Explorer. Summarizing tables for the output will be also produced.

Conscientious analysis of the outputs will allow to elaborate tactical scenarios for the selected areas of interest and taxa.

3.1.3 Input Data.

Forest areas populated with selected taxa will be analysed based on the National Forest Map. Strata for each pure species stands or existing mixed pull species will be delimited and prepared. Next step comprises the selection of plots from NFI associated to the each patch so all necessary plot data can be also prepare as input data for simulations.

Environmental variables, such as climatic, soil or topographic, are also needed to have accurate simulations. These data necessary came from other sources than Cross-Forest platform, but we need to attach them to our simulations.

3.1.3.1 Data preparation. Area by species selection and NFI data extraction.

Using the Spanish Forest Map, the complete list of patches (POLYGON in MFE dataset) for each species and admixtures composition will be selected (see Deliverable 3.2 annex I requirement 26-1). The list of necessary variables to identify areas of interest are described in Table 2.

Table 2 - Spanish Forest Map (MFE) variables in *Poligon* data-sheet.

Variable	Meaning / Values	Source / MFE name
Patch	Identifier of the area, unique for the province	(POLYGON)
Land_use	1 – Forest / 2 – Agriculture / 3 – Artificial / 4 – Wetland / 5 – Water	(nivel uso suelo 1) Filter to 1
Forest_use	1 – Forest / 2 – Open forest / 3 – Forest (temporarily clear) / 4 – Shrubs or pasture / 5 – Shrubs or pasture (Potentially forest) / 6 - Not continuous forest / 7 – Isolated trees	(nivel uso suelo 2) Filter to 1
Forested_use	1 – Forest / 2 – Plantations / 3 – Dehesa (agrosilvopastoral system) / 4 – Forest infrastructures	(nivel uso suelo 3)
Use	0 – Forested (Arbolado) / 1 – Shrubs (Matorral) / 2 – Grassland (Herbazal) / 3 – Non forested area (Monte sin vegetación superior) / 4 – Temporary unforested (Temporalmente desarbolado) / 5 – Forest infrastructures (Complementos del bosque) / 6 – Agriculture (Cultivo) / 7 – Artificial (Elementos artificiales) / 8 – Water (Agua) / 9 – Wetland (Humedales)	(NVal)
Species	Up to tree species presents in the patch	Esp1, Esp2, Esp3
Occupation	Proportion of the species, 0 to 10	O1, O2, O3
Stage (Stand age)	Stage of the species: young growth, sapling (replado), thicket (monte bravo), polewood (latizal), old growth, high forest (fustal), / low polewood	E1, E2, E3
ForestCover	Percentage of tree crown projection over forest area	Fcc, FccP, Fcctot

With this information, the NFI plots included in this patches will be selected using information from table “Parcpoly” which join NFI plot ID with patches ID. Further refining will be fulfilled by checking characteristics of minimum basal area, species correspondence and coherence in data, resulting in a first version of input data for simulations.

3.1.3.2 Simulator input data.

Input data for simulations can be provided in several formats, and although prior version of simulator SIMANFOR was fed with excel files, and this options keeps available, free open option are being implemented, such as JSON or CSV, making CROSS-FOREST Platform datasets (CFPd) accessible through SPARQL queries. Data variables are detailed in deliverable 3.2 (see Requirements 0 - 13 from Annex I), and it has tree main levels of concretion: (i) Invariant and related to study area, (ii) evolving with time and stand level attributes and (iii) time dependant tree level characteristics.

All variables describing study area are shown in Table 3, and are obtained from NFI and based on plot coordinates from CFPd, Digital Terrain Model (DTM) or “Catálogo de Montes de Utilidad Pública” (CMUP).

Table 3 - Study area variables

Variable	Meaning / Values	Source
Longitude and Latitude	as global position of the plot	CFPd from NFI – PCDatosMap.CoorX, CoorY
Altitude	Meter above sea level	DTM
Municipality	in which the plot is located	CFPd
Province	in which the plot is located	CFPd
StudyArea	Name of the geographical area in which the plot is locate	CFPd
Forest	identifier of property, number if public forest.	CMUP

Variable	Meaning / Values	Source
MainSpecies:	main species, up to tree, present in the plot	NFI – PCEspParc.Especie1, Especie2, Especie3
Plot_id	label to identify each plot	NFI – Cla, Subclase
Date	date of inventory	NFI – PCParcelas.FechaIni
Stage	level of development of the stand. young growth (Replado), thicket (Monte bravo), polewood (Latizal) or high forest (Fustal).	NFI – PCEspParc.Estado
StandType	Relationship of the age among trees. Even age, (Coetánea, Regular), two age classes (Semirregular) or uneven age (Irregular).	NFI - PCEspParc.FPMasa
InventoryAge	Age of the plot (Only valid when StandType is “even age”)	NFI - PCEspParc.Edad
SourceAge	Source of age value: 1- appearance / 2 -forest agent / 3 – forest service / 4 – locals / 5 – count of whorls / 6 – stups / 7 – other	NFI - PCEspParc.FInfor
ConfidenceAge	Confidence of age value: 1 – very low / 2 – low / 3 – medium / 4 – high	NFI - PCEspParc.Fiabil
StandSource_1	1 – Natural / 2 – Plantation / 3 – Previous plantation	NFI - PCEspParc.OrgMasa1
StandSource_2	1 – Seed / 2 – plantation / 3 – coppice / 4 – seed mixed with coppice / 5 – seed mixed with plantation / 6 – plantation mixed with coppice	NFI - PCEspParc.OrgMasa2
StandTreatment	1 – high or seedling forest / 2 – coppice with standards forest / 3 – coppice forest.	NFI - PCEspParc.TratMasa
SiteIndex	productivity of the area and species	calculus

A crucial variable for performing simulations is the age, if not possible at tree level at least at stand level, provided the stands we will use belong to the same class of age. In order to be able to estimate stand age we will also develop age equations relationships with dominant height and dominant height difference between inventories.

Variables describing stand characteristics should be calculated based on tree size variables and are detailed in Table 4.

Table 4 - Stand characteristics variables

Variable	Meaning / Values	Source
Density	Sum of expansion factor of all trees inside plot	NFI + calculus
QuadraticMeanDiameter	Square root of sum of square diameter of trees	NFI + calculus
BasalArea	Sum of sections of trees per hectare	NFI + calculus
DominantHeight	Average height of biggest 100 trees in the hectare	NFI + calculus
Age	Mean age for even-aged plots	NFI + calculus
Volume	Sum of volume of trees per hectare	NFI + calculus

Tree level variables are detailed in Table 5, for which initial value are provide by NFI and evolves with the simulation.

Table 5 - Tree level variables from NFI

Variable	Meaning / Values	Source
Inventory_id	as identification of the inventory	Unique Id. for the inventory

Variable	Meaning / Values	Source
Plot_id	as indentificator of the plot	Cla, Subclase
tree_id	as identification of the tree	PCMayores.nArbol, OrdenIf3, OrdenIf2
species	IFN code identification of tree species	PCMayores.Especie
dbh_1	Breast height diameter 1 in mm	PCMayores.Dn1
dbh_2	Breast height diameter 2 in mm	PCMayores.Dn2
quality	1 – healthy tree / 2 – few flaws / 3 – not healthy or dominated / 4 – weak tree, low quality wood / 5 – disease, old or very weak, very low quality wood / 6 – dead, low quality wood with monetary value	PCMayores.Calidad
shape	Shape which defines equation to calculate volume: 1 – fusiform stem, right, branches free in 6 m, dbh thicker than 20 cm / 2 - fusiform stem, right, branches free in 4 m / 3 - fusiform stem, smaller than 1 and 2 / 4 – branches below 4 m / 5 – stem bended, lot of branches or stem shorter than 4 m / 6 – pollarded trees of following species: oaks, poplars, beech, chestnut or bay.	PCMayores.Forma
height	Total height of the tree in m	PCMayores.Ht
dbh	Mean diameter in cm	PCMayores.(Dn1+Dn2)/20
expansion	Expansion factor to the hectare	calculate
vol	Volume of the tree over bark	VCC calculate
bole_vol	Volume of the tree under bark	VSC calculate
coord_x, coord_y	location of the tree, relative to the centre of the plot or absolute ones.	PCMayores.Rumbo, Distancia

For some growth models there are additional data requirements, such as climatic, soil, topographic, which should be provided to the simulator and for what connectors are being deployed in the environment.

3.1.4 Data transformation.

CAMBRIC pilot use data almost directly from Cross-Forest Platform as input for SIMANFOR simulator. All data transformations are related to the simulations itself and necessary calculus involved in scenarios and for obtaining dendrometric and dasometric variables, or the process for calculating the amount of wood of different sizes suitable for different industrial uses. Following are described all equations needed for each model.

3.1.4.1 Models Definition. Groups by species composition.

Models Definition. Mathematical models aiming to represent trees and stands evolution are defined, following the structure of IBERO model (Lizarralde et Al. 2010a and 2010b).

First we should select available growth and yield models for each taxa. Table 2 describes the different species and already fitted relationships, extracted from the analysis performed by Bravo et Al. (2012).

Necessary model equations are (see deliverable 3.2 Requirements 14 - 20 from Annex I):

- Site index equations (SI) for productivity assessment. It is usually given as the dominant height of the stand at a normalized age, often half of rotation age.

- Growth functions describing size changing during simulation period. Most frequent and useful is diameter (or basal area) growth functions (d.g), but height growth functions (h.g) can be applied to simulations.
- Mortality models (Mt) describes probability of dead for each size and stand conditions, which can be transform in a reduction of the value of variable “expan” in the inventory after a period of time.
- Ingrowth models (Ig) can predict the probability of new trees, which can be transform as increment in the value of the variable “expan” in the inventory after a period of time.
- Taper equations describes the shape of the stem of a tree. This function can be used to compute total volume and possible classification for different size pieces depending on its possible industrial destinies.
- Biomass (B) equations defining weight of dry material.
- Auxiliary equations which relates sizes of elements of the tree: bark thickness or bark percentage (bt), Height-diameter relationships (h-d), Crown equations (crw)...

Taxa for which equations availability will be checked are summarized in Table 6

Table 6: Taxa to be checked

Pure stands	Mixed stands	
Pinus sylvestris (Psy)	Psy + Ptr	Png + Qpy
Pinus pinaster (Ptr)	Psy + Png	Fsy + Qrb
Pinus pinaster atlantica (Ptr-A)	Psy + Fsy	Fsy + Qpt
Pinus nigra (Png)	Psy + Qpy	Qrb + Qpt
Fagus sylvatica (Fsy)	Psy + Qrb	Qrb + Qpy
Quercus pyrenaica (Qpy)	Psy + Qpt	Qpt + Qpy
Quercus robur (Qrb)	Ptr + Png	
Quercus petrea (Qpt)	Ptr + Qpy	

For species for which we did not find equations, we will develop new ones; For mixed stands we will also develop new models following Riofrío et Al. (2020) developed for Scotch’ and Mediterranean’ pine mixture, or whether this is not possible we will use the ones provided for mono-specific stands.

3.1.4.2 Industry Requirements. Classification of wood end use and quality

Industry Requirements. Raw material characteristics for each end use are defined, as well as the way to estimate the quantity available by year and area.

Taper equations included in each model will be used to calculate the different possible cutting schema for each bole, with optimum selection starting with biggest sizes.

The different industrial uses proposed for Cambric Pilot (see deliverable 3.2 Requirements 21.1 – 21.7 from Annex I) are the following, from small to large pieces of raw material needed:

- fuelwood, 8 to 20 cm in diameter and 1 m length,

- Round-wood:
 - timber purlins, rafters, stakes and sticks, 8 to 15 cm in diameter and 2.5 m length,
 - Post or pole-wood, 10 to 45 cm in diameter and 6 m length,
- Sawn-wood, 20 to 200 cm in diameter and 1.2 m length,
- Veneer, 15 to 160 cm in diameter and 2.6 m length,

Biomass estimation, total and for each component will be also provided. The equations developed by Ruiz-Peinado et Al. (2012 and 2013) will be used to achieve this result.

3.1.4.3 Simulation Definition. Silvicultural alternatives

Simulation Definition. Given a stand and trees characteristics at a year and age, a simulation shows how change and what silvicultural treatments are applied during a period of time. The five different silvicultural scenarios proposed by Duncker et Al. (2012) should be defined (see deliverable 3.2 Requirements 22 - 25 from Annex I), adapted to each species or species composition characteristics and possible industrial uses:

1. Passive scenario (Unmanaged Forest Nature Reserve), consisting in lack of management, that is, no cut planned along the life of the stand.
2. Low scenario (Close to nature forestry), consists in carrying out a management imitating the natural processes that occur in the forest, allowing exploitation of wood, although in small quantities, avoiding natural mortality. The operations must be compatible with the ecological functions of the forest.
3. Medium scenario (Combined objective forestry), two objectives are pursued, wood production and the achievement of other economic and ecological benefits.
4. High scenario (intensive even-aged forestry), whose main objective is the production of wood. Only *Pinus pinaster atlantica* taxa lays under the scope of this scenario (see Figure 2).
5. Intensive scenario (Short rotation forestry). This option will be only considered because for *Pinus pinaster* subsp *atlantica*, which is the sole species considered in the pilot suitable for this silviculture treatments.

3.1.5 Outputs

For each patch and strata all suitable simulations will be performed, aiming to obtain a detailed map of wood availability by industrial destiny and year (or period of time).

3.1.5.1 Simulation environment

The simulations will be performed with the new version of [SIMANFOR](#), a software solution flexible enough to easily adapt different equations for each pure or mixed species stands, now reprogrammed in Python, compatible with GNU Linux and able to run on HPC environments. This task has been externally performed by former developer of the simulator, [sngular](#).

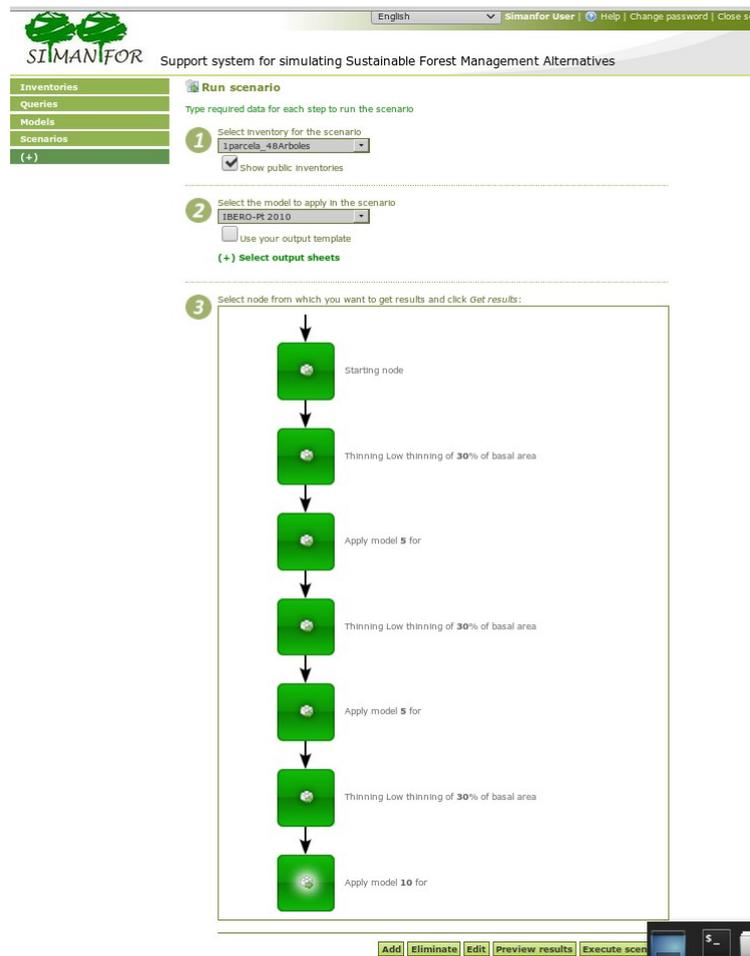


Figure 2 - High scenario for SIMANFOR

All simulations will start around the age of 25 years and end between 80 and 140 years except for *Pinus pinaster* subsp *atlantica* scenarios, which starts at 5 - 10 years and ends at 40 years, the most common rotation age for the species. We will avoid for all scenarios the regeneration period and early stage of the stand due to lack of suitable models and high complexity involved in this period of the stands.

3.1.5.2 Data output format

Output data will be of four types:

1. a table with the same information as input data but one by each node performed by the simulation;
2. a summary of the simulation with the starting point, the interval of years simulated, and the silvicultural operations performed (thinnings)
3. a summary report with the same information as included in a forest yield table
4. a production table for end uses previously proposed

Outputs type 2, 3 and 4 will be also included in Cross-Forest platform, so the ontology developed by GSIC/EMIC should be updated to include these data and make them accessible as LOD.

3.2 FRAME: Forest fiRes Advanced ModElization

3.2.1 Overall Perspective

In FRAME pilot, simulations of forest fire behavior will be carried out on real scenarios where fire will propagate over cell sizes of one square meter spatial resolution. Over these simulations, as a novelty respecting other pre-existing systems, real-time combat actions could be performed, as well as modifications of meteorological conditions, across a wide area of susceptible burning. Input values, which means digital elevation model (DEM), roads and infrastructures, fuel models and meteorological data, will be obtained from Open Data published by different public administrations.

Specifically, a modification of the propagation core developed in the wildfire virtual trainer ERVIN is used as the basis to develop FRAME propagation model, running over a HPC environment, and employing the emergencies and wildfire management system EMERCARTO (developed by Tragsa Group) as the user interface (see Deliverable 3.2 Annex IV).

Table 7 – FRAME pilot overall perspective

DESCRIPTION	Use of an HPC-adapted propagation model for predicting forest fire behavior and spreading depending on variables related to weather (wind, temperature and humidity), vegetation and terrain, using data from Portugal and Spain with high spatial and time resolution for a cross-border approach.	
PRECONDITION	The propagation model is running in an HPC environment. The databases of the Endpoint repository are accessible and published. The external databases are published/available. Access to the Endpoint repository and external databases is available.	
NORMAL SEQUENCE	STEP	ACTION
	1	Choose the ignition point, either by: <ul style="list-style-type: none"> - typing X,Y coordinates - drawing the point on a GIS interface (EMERCARTO).
	2	Enter a range of different weather conditions: wind, temperature and humidity. These conditions can be real or simulated.
	3	Run the HPC-adapted propagation model.
	4	Modify weather/vegetation conditions
	5	Run the HPC-adapted propagation model again and get a new iteration.

POSTCONDITION	<p>Quick mapping (and numerical information related) of fire behavior projections for the range of different weather conditions provided.</p> <p>Weather/vegetation conditions can be modified and new iterations are obtained.</p> <p>Geographical scope of fire behavior projections comprehends the Iberian Peninsula (Portugal and Spain).</p> <p>Publication of some of the results as Linked Open Data.</p>
----------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Figure 3 - ERVIN simulation screenshots

3.2.2 User interface

In this pilot, the resulting maps will be required to visualize in EMERCARTO, a GIS platform developed by Tragsa Group and employed for real-time emergency services monitoring, logistics and planning. EMERCARTO is employed by Tragsa Group and by several Public Administrations in Spain, and is accessible from either a PC or mobile devices, with different profiles and permissions.

Initial conditions and further modifications, as well as subsequent outputs, will be displayed in EMERCARTO system, where users will be able to interact with FRAME, in terms of changing weather conditions, locating ignition points, setting combat actions and receiving

final outcomes from calculations, such as perimeters of fire related to time intervals and fire statistics.



Figure 4 - EMERCARTO platform screenshot

3.2.3 Input data

a) From Cross-Forest LOD endpoint

Propagation models rely on information about fuel models. Information coming from Forest Inventories and Forest Maps, which will be accessible as LOD in the Cross-Forest Endpoint, will be the reference for building adapted fuel models that are efficient for running high resolution and simultaneous simulations over vast territories, according to variables operated by FRAME propagation core.

b) From external sources

To reach the main objective of FRAME, information about weather conditions and terrain with high spatial and time resolution will be used. The gathered data will allow the propagation model to run and to map fire behavior predictions based on real data. Therefore, it is indispensable that we identify the external databases, and the information we will need to use as input data.

To carry out the predictions and to present information in EMERCARTO, it will be necessary to have access to datasets of some public administration organizations (Portuguese and Spanish), which have databases about weather and updated digital elevation models (DEM). In Spain, AEMET and IGN will be key institutions as data providers. In Portugal, DGT and IPMA will provide the required data, as can be seen in the following table (see Deliverable 3.2 Annex V).

Table 8 – External datasets to be considered as input data for FRAME pilot

Themes	Variables	Data Sources	
		Spain	Portugal
Weather	Wind intensity and direction	AEMET WindNinja models	IPMA WindNinja models
	Temperature	AEMET	IPMA
	Humidity	AEMET	IPMA
Topography	Digital Elevation Model – (DEM)	IGN	DGT
Fuel	W, H, Vp, Tool Integrity	Spanish National Forestry Map	Land Cover National Map (COS) DGT
Roads and infrastructures	Individual defined cell assignment	IGN/Regional cartography services	DGT

3.2.4 Data transformation

According to the data specificities and their providers, the project team will define the best workflow process to access, employ and update the data if necessary.

This project will provide various tools to answer several questions, as showed in the following Table (see Deliverable 3.2 Annex VI).

Table 9 – FRAME pilot overview of questions to be answered

Questions
What will be fire perimeter at the hour "x"?
Will fire reach a specific point? At which hour?
What will be flame length or fire intensity at a specific point and hour?
What will be rate of spread at a specific point and hour?
Will fire exceed track or barrier, "x" meters wide, in his predicted progress?
Such a municipality will be affected? Such a natural space? In which proportion?

Which roads or paths will intersect the predicted progress of fire?
What will be the affected surface at a specific moment?
What will be maximum registered fire intensity and rate of spread? At what points?
Wind direction or slope at a certain place and at any given time?
What is the type of vegetation at a certain place and at any given time? What percentage of burned area, according to the type of vegetation, will be affected in each case?
Within an active perimeter at a given time, which is the most active point, that is, with higher rate of spread or fire intensity?

3.2.5 Outputs

The expected output of FRAME pilot is to present a series of maps showing fire behavior projections in several isochrones according to the storyline defined on an input stage, for the range of different weather conditions provided and with a geographical scope, which comprehends the Iberian Peninsula (Portugal and Spain).

Many iterations over the same initial points of ignition could be executed, varying, not only weather or fuel settings, but expected combat actions that will be performed on the wildfire scenario proposed.

In addition to fire perimeters displayed on EMERCARTO interface, this will provide tools to determine, on each cell of 1 square meter of computing, information about the main variables of fire behaviour, that is, rate of spreading, fire intensity, flame length and angle, residence time and fuel consumption, as well as information of input data obtained from scenario over that cell, such as type of fuel, slope or coordinates, for example.

These results will provide a useful management, prevention and training tool for Public Administrations and Emergency professionals, that nowadays utilise EMERCARTO services or those who could be interested in using it.

4 Conclusions

In this document, we present requirements for both pilots of Cross-Forest, CAMBRIC and FRAME. These requirements will be fulfilled by Cross-Forest platform, a Digital Service Infrastructure briefly presented in section 2 of the document, and composed (i) on the one side by a dataset published in Linked Open Data format, and (ii) on the other hand by a High Performance Computing infrastructure, Caléndula, which hosts the platform.

A separate description of each pilot and detailed specifications of data requirements are shown in section 3. All necessary changes accomplished in SIMANFOR and ERVIN simulators, aiming to run in Caléndula Supercomputer, have been also described. The description of each pilot helps to understand how data can be prepared to feed Cross-Forest platform and also to be used from external sources, and how important it is to have accessible datasets for performing the simulations planned in each pilot.

References

Bravo F, Álvarez-González JG, del Río M, Barrio M, Bonet JA, Bravo-Oviedo A, Calama R, Castedo-Dorado F, Crecente-Campo F, Condés S, et al., 2011. Growth and yield models in Spain: Historical overview, contemporary examples and perspectives. *Forest Syst* 20(2): 315-328. <http://dx.doi.org/10.5424/fs/2011202-11512>

Bravo, F.; Rodríguez, F.; Ordóñez, A.C. A web-based application to simulate alternatives for sustainable forest management: SIMANFOR. *For. Syst.* 2012, 21, 4–8.

Dunker, P., Barreiro, S., Hengevelds, G., Lind, T., Mason, W., Ambrozy, S., & Spiecker, H. (2012). Classification of forest management approaches: A new conceptual framework and its applicability to European forestry. *Ecology and Society*, 17(4), 51.

Gonzalvo, V.; Aguirre, F.; Carrillo, A.; López, J.; González, D.; Díaz, M.; Benito, I. Tragsa Group and wildfires. AF3 Project. Technology integration on fires. *Revista Montes* (ISSN 0027-0105), number 127, Madrid, 2017.

Gonzalvo, V.; Hombrados, E.; Aguirre, F.; Carrillo, A. ERVIN instructs virtually in wildfires. *Revista Montes* (ISSN 0027-0105), number 130, Madrid, 2017.

Gonzalvo, V.; Hombrados, E.; Aguirre, F.; Carrillo, A. ERVIN: A virtual trainer for wildfires. Spanish Forest Conference 2017, <http://7cfe.congresoforestal.es/content/ervin-un-entrenador-virtual-para-los-incendios-forestales>, Plasencia, 2017.

Gonzalvo, V.; Aguirre, F.; Carrillo, A.; López, J.; Díaz Romero, M., Benito, I. AF3: Integration of assistive technologies in fire management. Spanish Forest Conference 2017, <http://7cfe.congresoforestal.es/content/af3-integracion-de-tecnologias-de-apoyo-en-la-gestion-de-incendios>, Plasencia, 2017.

Heym, M., Ruíz-Peinado, R., Del Río, M., Bielak, K., Forrester, D. I., Dirnberger, G., ... Pretzsch, H. (2017). EuMIXFOR empirical forest mensuration and ring width data from pure and mixed stands of Scots pine (*Pinus sylvestris* L.) and European beech (*Fagus sylvatica* L.) through Europe. *Annals of Forest Science*, 74(3). <https://doi.org/10.1007/s13595-017-0660-z>

Lizarralde I, Ordóñez AC, Bravo-Oviedo A, Bravo F. 2010a. IBERO.PT: Modelo de dinámica de rodales de *Pinus pinaster* Ait. en el sistema ibérico meridional (available online at www.simanfor.org).

Lizarralde I, Ordóñez AC, Bravo-Oviedo A, Bravo F. 2010b. IBERO.PS: Modelo de dinámica de rodales de *Pinus sylvestris* L. en el Sistema Central y el Sistema Ibérico en Castilla y León (available online at www.simanfor.org).

Pretzsch, H., & Schütze, G. (2014). Size-structure dynamics of mixed versus pure forest

stands. *Forest Systems*, 23(3), 560–572. <https://doi.org/10.5424/fs/2014233-06112>

Riofrío, Bravo F, Rio M (2019) Adapting the tree-level growth and yield model IBERO to mixed species stands (manuscript)

Rodríguez, F., Lizarralde, I., & Bravo, F. (2015). Comparison of stem taper equations for eight major tree species in the Spanish Plateau. *Forest Systems*, 24(3). <https://doi.org/10.5424/fs/2015243-06229>

Ruiz-Peinado R, Rio M, Montero G (2011) New models for estimating the carbon sink capacity of Spanish softwood species. *Forest Syst* 20:176–188. doi: [10.5424/fs/2011201-11643](https://doi.org/10.5424/fs/2011201-11643)

Ruiz-Peinado R, Montero G, Rio M (2012) Biomass models to estimate carbon stocks for hardwood tree species. *Forest Syst* 21:42–52. doi: [10.5424/fs/2112211-02193](https://doi.org/10.5424/fs/2112211-02193)

The Open MPI Project ©2004-2020. MPI - General information Open MPI 4.0.2. URL <https://www.open-mpi.org>

Cross-Forest proposal. (2017). Action Number 2017-EU-IA-0140.

Cross-Forest Grant Agreement. Agreement Number: INEA/CEF/ICT/A2017/1566738

Cross-Forest “D1.1 Operational analysis and specifications. v.2”

Cross-Forest “D2.1 Data collection, Data quality, Ontologies and Vocabularies”

Cross-Forest “D2.2 Data exportation and Publication - Interim Report”

Cross-Forest “D3.2 Cross-Forest pilots. First iteration”

Cross-Forest “D4.3 Evaluation of results - Interim Report”

Definitions, acronyms and abbreviations

AEMETSpanish Meteorological Agency
CALENDULASupercomputer located at SCAYLE
CAMBRICCALidad de la Madera en Bosques mlxtos (Wood quality in mixed forest)
CBDCrown Bulk Density
CEFConnecting Europe Facility
CFISCrown fire initiation and spread
CORDISCommunity Research and Development Information Service
DEMDigital Elevation Model
DGTDireção-Geral do Território
DTMDigital Terrain Model
ECEuropean Commission
EMERCARTOEmergencies and wildfire management system
ERVINVirtual Trainer for Wildfires
EUEuropean Union
FRAMEForest fiRes Advanced ModElization
GISGeographic information system
HPCHigh performance Computing
ICTInformation and communications technology
ICNFInstituto de Conservação da Natureza e Florestas
IGNNational Geographic Institute (Spain)
INEAInnovation and Networks Executive Agency
IPMAPortuguese Institute for the Sea and Atmosphere
LODLinked Open Data
MAPAAgriculture and Environment Ministry (Spain)
NFINational Forest Inventory
SCAYLEFundación Centro de Supercomputación Castilla y León (Castilla y León Supercomputing Center)
SIMANFORSupport system for simulating Sustainable Forest Management Alternatives
TRAGSAEmpresa de Transformación Agraria, S.A.
UCOUniversidad de Córdoba
UVA (IuFOR)University of Valladolid (Institute for Research in Sustainable Forest Management)
UVA (GSIC/EMIC)University of Valladolid (Group of Intelligent & Cooperative Systems / Education, Media, Informatics & Culture)