Ref. Ares(2019)3623741 - 05/06/2019

Horizon 2020 Societal challenge 5: Climate action, environment, resource efficiency and raw materials

VERIFY
Observation-based system for monitoring and verification of greenhouse gases
GA number 776810, RIA

Deliverable number (relative in WP)	D3.1
Deliverable name:	First state of the art database
WP / WP number:	3
Delivery due date:	Month 14 (March 2019)
Actual date of	Month 16 (31/05/2019)
Dissemination level:	Public
Lead beneficiary:	University of Aberdeen
Responsible	Matthias Kuhnert
Contributor(s):	Matthias Kuhnert, Pete Smith, Matthew McGrath, Philippe Peylin, Karina Winkler, Richard Fuchs, Martin Herold, Emanuele Lugato, Richard Engelen, Adrian Leip, Philippe Ciais, Mart-Jan Schelhass, Gert-Jan Nabuurs, Frank Deneter, Ronny Lauerwald, Pierre
Internal reviewer:	Matthew McGrath

1. **Changes with respect to the DoA**

None

2. **Dissemination and uptake**

The collected data are free available (in some case registration is/will be necessary). In the project these data build the basis for model simulations in WP3 and WP4. The web-pages for data download and or description is listed in section 3. The provided data build the basis for running the models as well as for calibration and validation of the bottom up ecosystem models.

3. **Short Summary of results (<250 words)**

For the quantification of GHG emissions models play a crucial role. They can extrapolate and interpolate measurements spatially and temporally. The application of models requires data sets to parameterize the models and validate the results. Climate, soil, management (for cropland, forest and grassland) and land use/land cover are the main driving data that are required for modelling of the carbon and nitrogen dynamics.
Therefore, the aim of this WP is to collect these data and provide them to the relevant work packages (WP3 and WP4). Additionally, the collection includes data to improve the simulations (N2 deposition) and to validate (Flux data sets) the result. Additionally, coastal ocean and fresh water fluxes are essential to complete the budget for the top down approaches in WP3 and WP4.
This deliverable provides details about the actual status of the database. Data for spatial simulation are collected and summarized in table 1. For almost all tasks at least preliminary data are provided. Some tasks are still working on improvements and extensions of the actual data sets and subsequent updates of this deliverable will be provided during the course of the project.

4. **Evidence of accomplishment**

All the dataset will be accessible though the VERIFY web site and the dedicated data-products page:
http://verify.lsce.ipsl.fr/index.php/products
Note that some of these data may be password protected during a consolidation phase and thus only accessible to the VERIFY partners (accessible through the internal share-point platform). Most data have been uploaded on the portal and are accessible through a catalogue and a thredds server (see section 3).

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

Version	Date	Description	Author (Organisation)
V0	27/05/19	Creation/Writing	Matthias Kuhnert (University of Aberdeen)
V1	04/06/19	Writing/Formatting/Delivery	Matthias Kuhnert, Pete Smith, Matthew McGrath, Philippe Peylin, Karina Winkler, Richard Fuchs, Martin Herold, Emanuele Lugato, Adrian Leip, Philippe Ciais, Mart-Jan Schelhass, Gert-Jan Nabuurs, Frank Deneter, Ronny Lauerwald, Pierre Regnier, Are Olsen, Maximilian Reuter
V2	05/06/19	Formatting/Delivery on the Participant Portal	Matthias Kuhnert, Pete Smith, Matthew McGrath, Philippe Peylin

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

1. Introduction

We describe below a first version of the datasets that will be used throughout the VERIFY project either as input to the different ecosystem models (both for WP3 and WP4) or for the evaluation of the simulated greenhouse gas fluxes and carbon stocks. A summary of the different dataset is provided in the table below.

Table 1.1: Available data set in the context of VERIFY: from the internet (CORINE and HWSD) and provided by project participants.

Data set	Name/ model	Inst.	Coverage	Resolution	Time frame	Contact in the project
Land use	CORINE[1]	EEA	global	0.1 km	2000, 2006, 2012	
Soil	HWSD[2]	FAO	global	1 km		
Land use	HILDA[3]	KIT	global	1 km	1900-2010	Richard Fuchs[a], Karina Winkler[b]
Biomass	HILDA[3]	WU	global	1 km	1900-2010	Martin Herold[c]
N-Deposition	EMEP model[4]	JRC		1 km	2010	Frank Dentener[d]
Erosion	PESERA[5]	ESDAC	Europe	1 km	2004	Emanuel Lugato[e]
Soil data	LUCAS[6]	ESDAC	Europe	1 km/10 km	2015	Emanuel Lugato[e]
Climate	C3S-ERA-5[7]	ECMWF	global	31 km	2008-current	Richard Engelen[f]
Climate	UERRA[8]	ECMWF	Europe	11 km	1961-2017	Richard Engelen[f]
Flux data	FLUXNET[9] network		global	sites	diverse	Dario Papale, Werner Kutsch
Fertiliser application rates	CAPRI	JRC	Europe	0.25°	2000-2012	Adrain Leip[g]
Crop management timing		UNIABDN	EU28	0.25°	2000-2015	Matthias Kuhnert[h]
Fresh water fluxes		ULB	Europe	0.1°	2016	Ronny Lauerwald[i]
Ocean coastal fluxes		UiB	Northern Europe	0.125°	1997-2016	Are Olsen[j]
Forest management		WU	Europe	0.125°	2000-2015	Mart-Jan Schelhass[k], Gert-Jan Nabuurs[l]
Grassland		CEA-LSCE	Europe	0.5°	1860-	Philippe Ciais[m]

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

management					2012	
atmospheric CO2	FOCAL[10]	UBremen	global	2 km	2015-2016	Maximilian Reuter[n]

1 https://www.eea.europa.eu/publications/COR0-landcover
2 http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/
3 https://www.wur.nl/en/Research-Results/Chair-groups/Environmental-Sciences/Laboratory-of-Geo-information-Science-and-Remote-Sensing/Models/Hilda/HILDA-data-downloads.htm
4 http://webdab.emep.int/Unified_Model_Results/
5 https://esdac.jrc.ec.europa.eu/content/pan-european-soil-erosion-risk-assessment-pesera
6 https://esdac.jrc.ec.europa.eu/resource-type/european-soil-database-soil-properties
7 https://www.ecmwf.int/en/about/media-centre/science-blog/2017/era5-new-reanalysis-weather-and-climate-data
8 https://confluence.ecmwf.int//display/UER
9 https://fluxnet.fluxdata.org/login/?redirect_to=/data/download-data/
10 http://www.iup.uni-bremen.de/~mreuter/TN_XCO2-OCO2-FOCAL_v08.pdf

Email addresses of contact persons:

a richard.fuchs@kit.edu
b karina.winkler@kit.edu
c martin.herold@wur.nl
d frank.dentener@ec.europa.eu
e Emanuele.LUGATO@ec.europa.eu
f richard.engelen@ecmwf.int
g Adrian.LEIP@ec.europa.eu
h Matthias.kuhnert@abdn.ac.uk
i rlauerwa@ulb.ac.be
j are.olsen@uib.no
k martjan.schelhaas@wur.nl
l gert-jan.nabuurs@wur.nl
m philippe.ciais@lsce.ipsl.fr
n mreuter@iup.physik.uni-bremen.de

2. Description of the different input and forcing datasets for WP3 activities

The different products that are described below follow the grouping proposed in the Document of Work (DoW).

2.1. State of the art climate data

The 55 year reanalysis run was performed using the HARMONIE system cycle 38h1.1. HARMONIE is basically a script framework that allows for different physics packages, surface schemes or data assimilation schemes. In the long UERRA run several changes in the script system were made, compared to the reference version of HARMONIE, to speed up the code. The main achievement was to separate the analysis and forecast steps. In the UERRA runs the new analysis is started as soon as the first guess is available, i.e. the 6 hour forecast. The remaining forecast hours is run in parallel to the next analysis. This saves a lot of time in a reanalysis but is of no use for operational forecasts. The ALADIN synoptic scale physics scheme was used together with a three dimensional variational data assimilation (3D-Var) scheme including only conventional observations and an OI assimilation scheme for the surface observations.

The original data for 1961-2018 was downloaded from ECMWF at hourly temporal resolution, and then regridded to convert from the Lambert projection to a rectilinear grid covering 35N:73N, 25W:45E, at 0.125 degree resolution (305x561). Remapping is bilinear except for precipitation variables and wind direction, which use nearest neighbour. From the raw data, the rainfall rate and the north and east components of the wind direction were computed, as was the maximum and minimum temperature, giving a total of 14 variables: Tair, Tmax, Tmin, Wind_N, Wind_E, WS, Psurf, LWdown, SWdown, Qair, RH, Rainf, Snowf, and Precipitation. Different groups require different temporal resolutions for their models: sub-daily, daily, and monthly. Files were produced based on the needs of WP3 gathered through a survey.

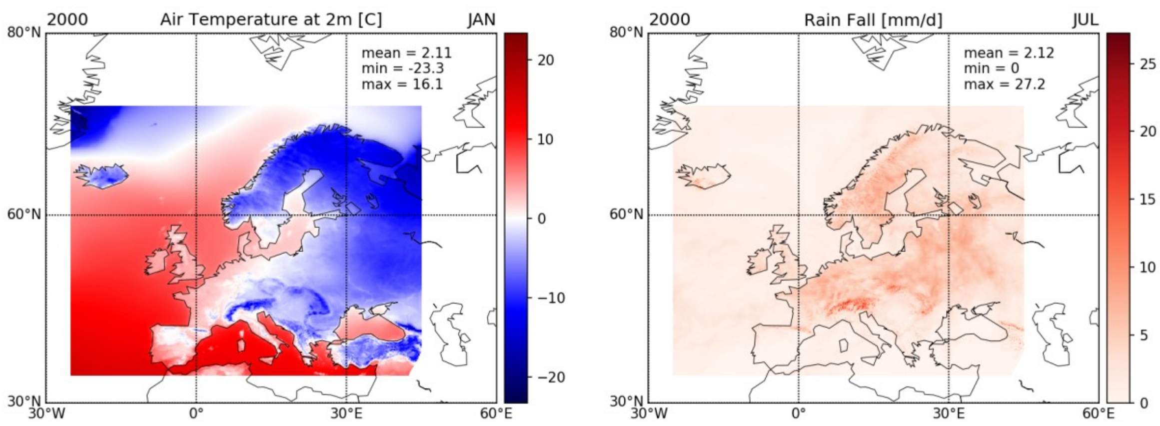


Figure 2.1.1: Illustration of climate forcing that will be used in VERIFY to force ecosystem models: left: surface air temperature for January 2000; right: rainfall in July 2000.

2.2. Land use datasets and high resolution land cover change and biomass mapping

2.2.1. High resolution land cover

We provide a European subset (beta version) of the HILDA+, a global data set on land use/land cover (LULC) change. We developed the HILDA+ land use/land cover maps using a data-driven reconstruction approach.

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

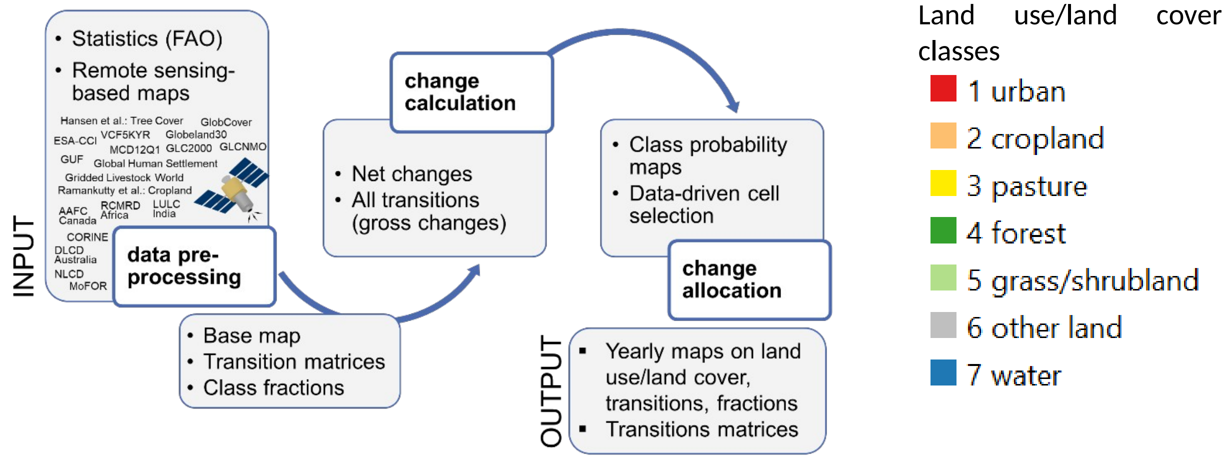


Figure 2.2.1.1. Data-driven approach for land change allocation

Starting with a FAO-calibrated base map (derived from ESA Copernicus LC100 2015), we allocated land use/land cover transitions iteratively for each time step (yearly) and for each country along a backward-looking time loop on a 1x1 km grid. Net change magnitudes are based on national FAO land use and population statistics. Gross change magnitudes are calculated from mean transition matrices, which are extracted from time series of satellite-derived land use/land cover maps. The change allocation depends on class probability maps (mean class fractions) generated from year- and region-specific remote sensing-based land use/land cover maps (see Table 2.2.1.1).
Multiple observational data sets (national inventories and earth observation-based maps) are used as input data source for the reconstruction process:

Table 2.2.1.1. Input data and references used for HILDA+

Dataset and reference	Thematic coverage	Spatial coverage	Used Temporal coverage	Spatial resolution	Data type
FAOSTAT http://www.fao.org/faostat/	Land use • Cropland and arable land • Pastures and meadows • Forest Population	global	1961-2015 1990-2015 1950-2015	national level	table
GLAD UMD Global Land Cover Change VCF https://glad.umd.edu/dataset/long-term-global-land-change	tree canopy bare ground	global	1982-2015	0.05 degree	raster

	short vegetation				
ESA Copernicus LC100 https://land.copernicus.eu/global/products/lc	LCCS 22 classes	global	2015	100 m	raster
ESA CCI Land Cover http://maps.elie.ucl.ac.be/CCI/viewer/download.php	LCCS 22 classes	global	1992-2015	300 m	raster
Globeland30 http://www.globeland30.org	10 LULC classes	global	2000, 2010	30 m	raster
Global Human Settlement Layer (GHSL) https://ghslsys.jrc.ec.europa.eu/datasets.php	Built-up area (fractional)	global	1975, 1990, 2000, 2014	1 km	raster
Global Urban Footprint (GUF) https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-11725/20508_read-47944/	Built-up area (fractional)	global	2011/12	2.8 arc sec	raster
Hansen tree cover https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.5.html	Tree cover (fractional) Loss and gain year	global	2000-2015	30 m	raster
Ramankutty cropland and pastures http://www.ramankuttylab.com/data.html	Cropland pastures	global	2000	5 arc min	raster
Gridded Livestock World v3 (GLW) http://www.fao.org/livestock-systems/en/	Density of ruminants	global	2010	5 arc min	raster
GLC2000 http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php	FAO LCCS 22 classes	global	2000	1 km	raster
MODIS Land Cover MCD12Q1 https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1	IGBP 17 classes,	global	2001-2013 (yearly)	500 m	raster
GlobCover http://due.esrin.esa.int/page_globcover.php	LCCS 22 classes	global	2005/2006, 2009	300 m	raster
GLCNMO https://globalmaps.github.io/glcnmo.html	LCCS 22 classes	global	2003 2008, 2013	1 km 500 m	raster
LULC classification of India (Meiyappan et al., 2017) https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1336	11 LULC classes (IGBP scheme)	national India	1985, 1995, 2005	100 m	raster
RCMRD	6 LULC classes with country-specific sub-classes	Regional: Botswana, Ethiopia, Lesotho.	different years between 2000 and 2014	30 m	raster

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

		Malawi, Namibia, Rwanda, Tanzania, Uganda, Zambia			
CORINE https://land.copernicus.eu/pan-european/corine-land-cover/view	44 LULC classes with change layers	regional / continental: Europe	1990, 2000, 2006, 2012	100 m	raster
NLCD https://www.mrlc.gov/finddata.php	16 LULC classes	national USA	2001, 2006, 2011	30 m	raster
Australia Dynamic Land Cover DLCD V2.1 https://ecat.ga.gov.au/geonetwork/srv/eng/catalog.search#/metadata/83868	LCCS 22 classes	national Australia	2002-2014	500 m	raster
Canada https://open.canada.ca/data/en/dataset/18e3ef1a-497c-40c6-8326-aac1a34a0dec	15 LULC classes	national Canada	1990, 2000, 2010	30 m	raster
MoFOR Indonesia http://webgis.dephut.go.id:8080/kemenhut/index.php/en/feature/download	22 LULC classes	national Indonesia	2000, 2003, 2006, 2009	300 m	raster
South Africa Land Cover	35/72 LULC classes	National South Africa	1990, 2013-14	30 m	raster

For the time when no observational data sets are available, remote sensing products with a sufficiently long time series (ESA CCI, MODIS MCD12Q1, GLAD UMD VCF) were back-casted/extrapolated based on the mean trend of the first five observed values in time.

The delivered European subset (beta version) of the HILDA+ land use/land cover data set has the following general specifications:

Table 2.2.1.2. Specifications of HILDA+

Spatial resolution	*1 km x 1 km*
Spatial coverage	Continental Europe (subset; see); global (entire data set)
Temporal resolution	1 year (1960-2015); 5 years (1900-1960) for backward extension
Temporal coverage	1960-2015; 1900-1960 for backward extension
LULC classes	• Urban (built-up areas, settlement, infrastructure, mining)

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

	• Cropland (agricultural land used for crop cultivation) • Pastures (managed grassland for grazing/livestock) • Forest (forested areas according to FAO definitions, including tree plantations) • Grass/Shrubland (grass- or shrubland areas under no or little management) • Other land (non-vegetated land: snow/ice, barren, rocks)
Data format and content	NetCDF files • Land use/land cover states for each year (coded with each LULC class number at each time step) • Land use/land cover transitions for each time step (coded with class transitions number between time steps T0 and T1) CSV tables • Land transition matrices for each country and each time

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

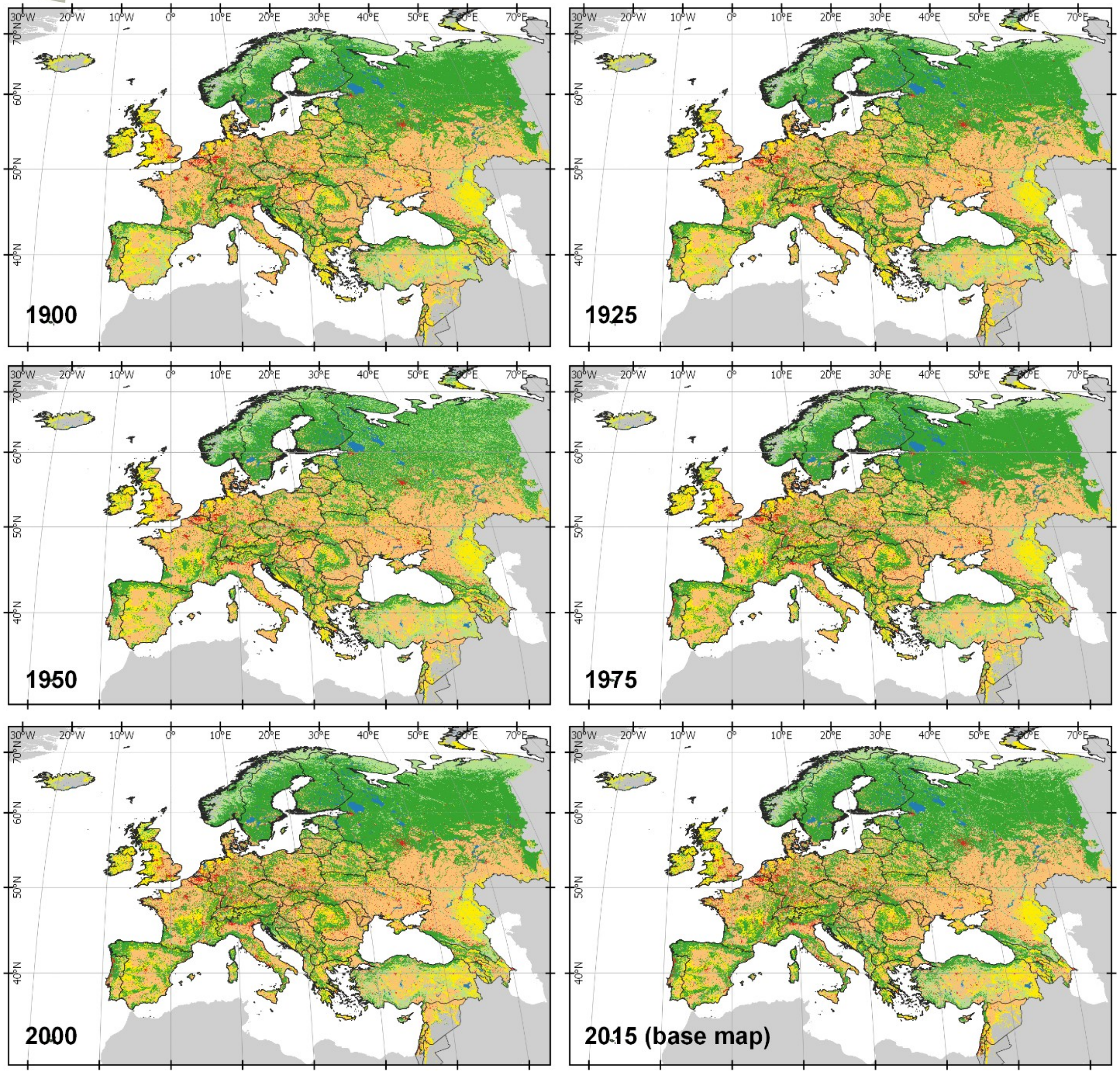


Figure 2.2.1.1. HILDA+ land use/land cover maps for selected years

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

2.2.2. Biomass maps

The biomass maps will be build up on the land cover maps that are provided by task 2.2.1. The approach will improve the spatial assessment of aboveground biomass by combining field observations, remote sensing and auxiliary data, from local case studies to global scale. The approach will build up on the GEOCARBON products (www.wageningenur.nl/grsbiomass) and the ESA-BIOMASS mission. This work is still in progress as it builds up on the land cover data.

2.3. Soil property and soil erosion datasets

2.3.1 Topsoil physical properties for Europe (based on LUCAS topsoil data)

The Land Use and Cover Area frame Statistical survey (LUCAS) aimed at the collecting harmonized data about the state of land use/cover over the extent of European Union (EU). Among these 200k land use/cover observations selected for validation, a topsoil survey was conducted at about 10% of these sites. Topsoil sampling locations were selected as to be representative of European landscape using a Latin hypercube stratified random sampling, taking into account CORINE land cover 2000, the Shuttle Radar Topography Mission (SRTM) DEM and its derived slope, aspect and curvature.

The LUCAS topsoil database was used to map soil properties at continental scale over the geographical extent of Europe. Several soil properties were predicted using hybrid approaches like regression kriging. For those datasets, topsoil texture and related derived physical properties were predicted. Regression models were fitted using, along other variables, remotely sensed data coming from the MODIS sensor. The high temporal resolution of MODIS allowed detecting changes in the vegetative response due to soil properties, which can then be used to map soil features distribution. The prediction of intrinsically collinear variables like soil texture required the use of models capable of dealing with multivariate constrained dependent variables like Multivariate Adaptive Regression Splines (MARS). Cross validation of the fitted models proved that the LUCAS dataset constitutes a good sample for mapping purposes leading to cross-validation R2 between 0.47 and 0.50 for soil texture and normalized errors between 4 and 10%.

This dataset provides the following soil properties at 500 m resolution, for the geographical coverage: European Union (EU) plus Balkan countries, Switzerland and Norway.

- Clay content (%) in topsoil (0-20cm) modelled by Multivariate Additive Regression Splines
- Silt content (%) in topsoil modelled by Multivariate Additive Regression Splines
- Sand content (%) in topsoil modelled by Multivariate Additive Regression Splines
- Coarse fragments (%) content in topsoil modelled by Multivariate Additive Regression Splines
- Bulk density derived from soil texture datasets (obtained from the packing density and the mapped clay content following the equation of Jones et al. 2003)

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

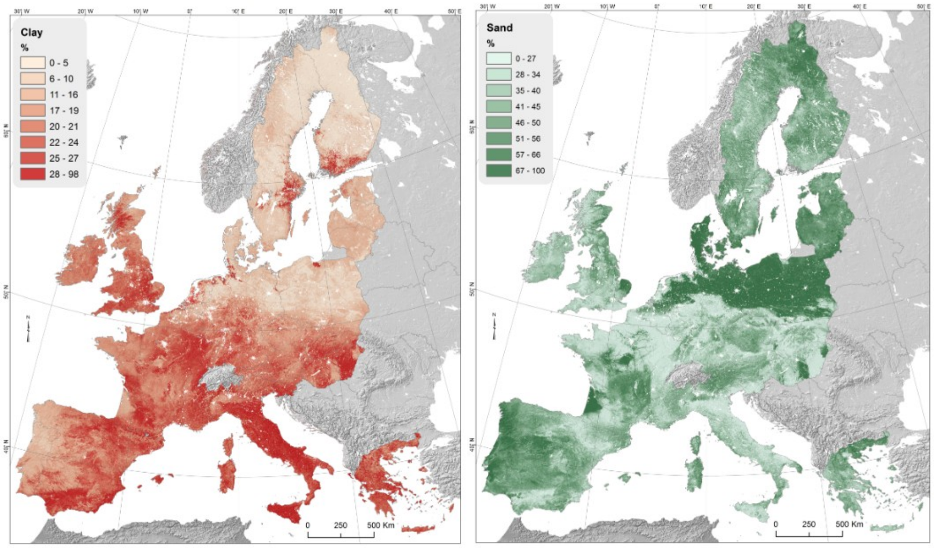


Figure 2.3.1.1 – Example of spatial layers provided (clay and sand content)

References:

Ballabio C., Panagos P., Montanarella L. Mapping topsoil physical properties at European scale using the LUCAS database (2016) Geoderma, 261 , pp. 110-123.
https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data

2.3.2.Soil erosion by water (RUSLE2015)

The new soil loss by water erosion map of Europe uses a modified version of the RUSLE model (RUSLE2015, based on Renard et al., 1997), which calculates mean annual soil loss rates by sheet and rill erosion according to the following equation:

$E = R * K * C * LS * P$

Where

E: Annual average soil loss (t ha^{-1} yr^{-1}),
R: Rainfall Erosivity factor (MJ mm ha^{-1} h-1 yr^{-1}),
K: Soil Erodibility factor (t ha h ha^{-1} MJ^{-1} mm^{-1}),
C: Cover-Management factor (dimensionless),
LS: Slope Length and Slope Steepness factor (dimensionless),
P: Support practices factor (dimensionless).

The RUSLE2015 improves the quality of estimation by introducing updated (2010), high-resolution (100 m), peer-reviewed input layers:

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

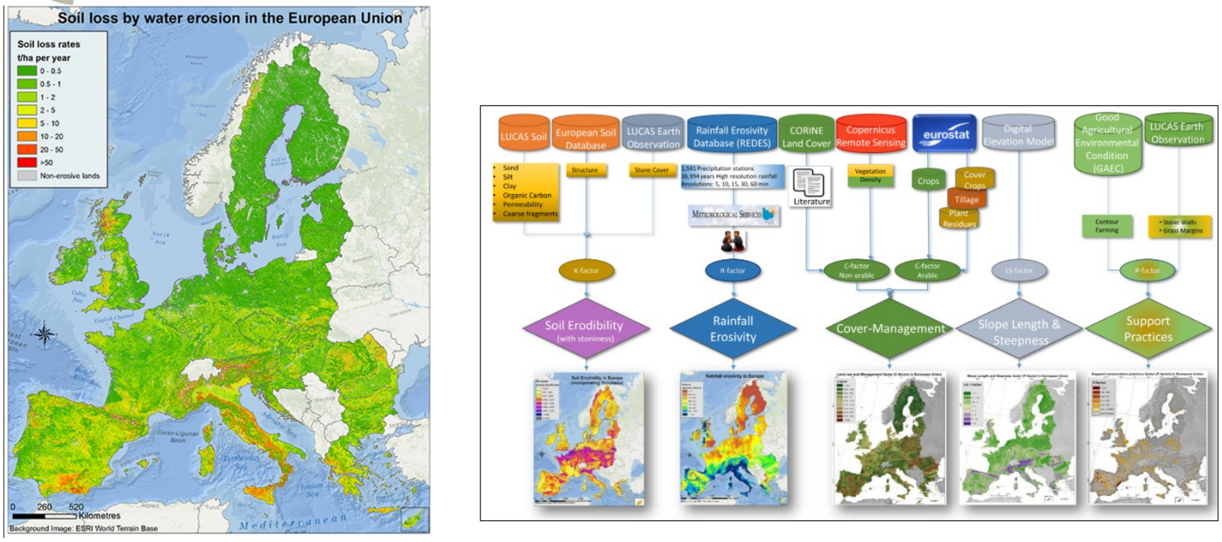


Figure 2.3.2.1 – Soil loss by water erosion (t ha^{-1} yr^{-1}) in the EU and RUSLE input framework.

References:

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., Alewell, .C. 2015. The new assessment of soil loss by water erosion in Europe. Environmental Science & Policy. 54: 438-447. DOI: 10.1016/j.envsci.2015.08.01
https://esdac.jrc.ec.europa.eu/themes/rusle2015

2.3.3.Soil erosion by wind (GIS-REWQ)

GIS-RWEQ is a simplified GIS-based application of the Revised Wind Erosion Equation-RWEQ model (ARS-USDA). It follows a spatially distributed approach based on a grid structure, running in R and Python scripts. The model scheme is designed to describe the daily soil loss potential at regional or larger scale. This dataset provides the annual average soil losses (t ha^{-1} yr^{-1}) by wind at 1km resolution for 28 Member States of the European. More specifically, the study area covered the following CORINE 2006 land cover unit: non-irrigated arable land (code 2.1.1) and permanently irrigated land (code 2.1.2). The resulting modelling area amounted to ca. 96.1 million hectares.

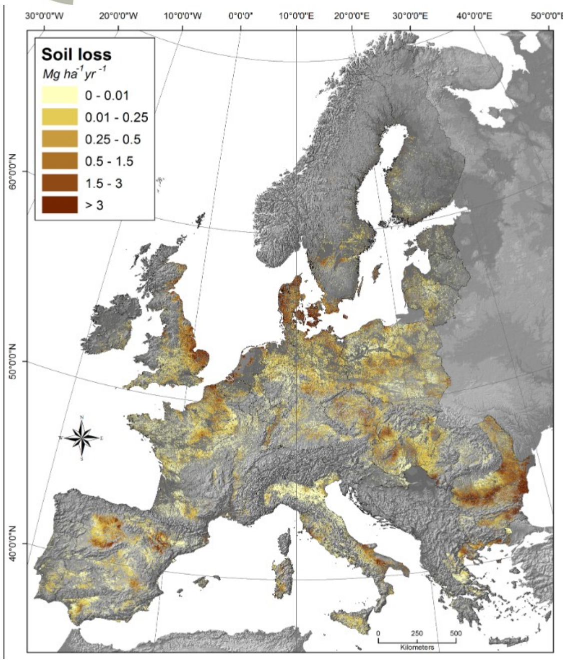


Figure 2.3.3.1. Soil loss by wind erosion in European agricultural soils

References:

Borrelli, P., Lugato, E., Montanarella, L., & Panagos, P. (2017). A New Assessment of Soil Loss Due to Wind Erosion in European Agricultural Soils Using a Quantitative Spatially Distributed Modelling Approach. Land Degradation & Development, 28: 335-344, DOI: 10.1002/ldr.2588
https://esdac.jrc.ec.europa.eu/themes/wind-erosion-european-agricultural-soils-rweq

2.3.4. Eroded soil organic carbon

This spatial layer derived from a consistent biogeochemistry-erosion model framework able to dynamically quantify the lateral C fluxes, driven by water erosion. The framework couples the process-based biogeochemistry model CENTURY and the RUSLE2015 erosion model and runs at 1km^2 resolution in the agricultural soils of the EU.

In addition, the model integrates both the erosion and depositional processes at pixels level in parsimonious way, calculating the following C budgets:

For the eroding area the C balance is:

dSOC = CI + CS – CH - CE - DOC

where dSOC is the SOC stock change in the fixed profile, CI is the C input through remaining NPP (after C exportation by harvest, but including roots and manure), CH is the heterotrophic respiration, CS is the incoming SOC from deeper layer, CE is the lateral C flux by sediment transport and DOC the C exported as dissolved organic carbon.

For the depositional areas the balance is:

dSOC = CI + CD - CH - CB - DOC

where CD is the C deposited coming from eroding areas and CB is the C that is moved out (i.e., buried) of the simulated profile as a consequence of soil deposition.

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

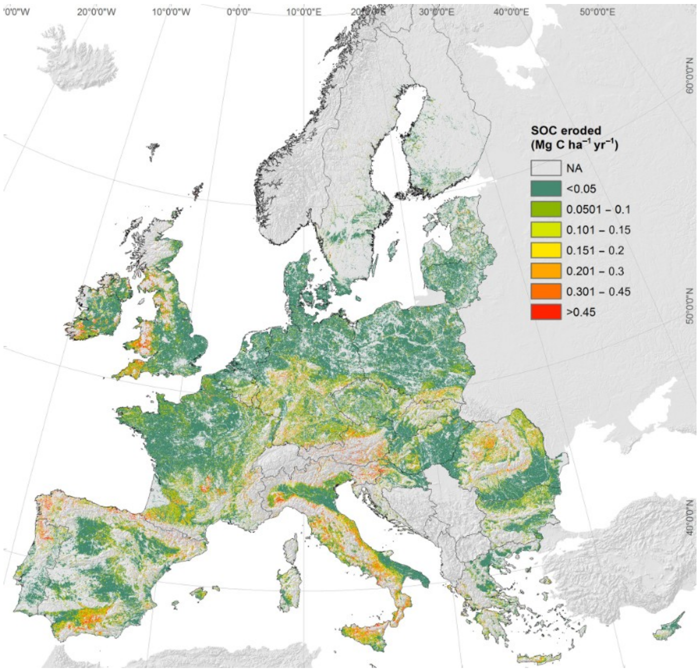


Figure 2.3.4.1. Average gross eroded SOC (Mg C ha^{-1} y^{-1}) in the period 2010-2100.

References:

Lugato, E., Smith, P., Borrelli, P., Panagos, P., Ballabio, C., Orgiazzi, A., Fernandez-Ugalde, O., Montanarella, L., Jones, A. Soil erosion is unlikely to drive a future carbon sink in Europe (2018) Science Advances, 4 (11), art. no. eaau3523.

Lugato, E., Paustian, K., Panagos, P., Jones, A., Borrelli, P. Quantifying the erosion effect on current carbon budget of European agricultural soils at high spatial resolution (2016) Global Change Biology, 22 (5), pp. 1976-1984.

https://esdac.jrc.ec.europa.eu/content/pan-european-soc-stock-agricultural-soils

https://esdac.jrc.ec.europa.eu/themes/soil-erosion-and-carbon

2.3.5. Soil Hydraulic Database of Europe at 1 km and 250 m resolution

The multilayered European Soil Hydraulic Database (EU-SoilHydroGrids ver1.0) was derived with European pedotransfer functions (EU-PTFs; Tóth et al., 2015) based on the soil information of SoilGrids250m and aggregated 1 km datasets.

It covers the parameters:

- saturated water content,
- water content at field capacity and wilting point,
- saturated hydraulic conductivity and Mualem-van Genuchten parameters for the description of the moisture retention
- unsaturated hydraulic conductivity curves
-

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

The EU-PTFs (Tóth et al., 2015) were trained on the European Hydropedological Dataset (EU-HYDI; Weynants et al., 2013). EU-HYDI is a collection of data from 29 institutions in 18 European countries and contains data on taxonomical, chemical, and physical soil properties of more than 18,000 soil samples. Pedotransfer functions were calibrated using soil information of 134 to 6,074 soil samples and validated on 57 to 2,357 samples, depending on the type of soil hydraulic property (Tóth et al., 2015).
SoilGrids provides the most detailed information on soil properties with full continental coverage in Europe. It incorporates soil taxonomical, physical, and chemical data of seven soil depths at 250 m resolution. The following soil properties to calculate the soil hydraulic properties were used: clay, silt, and sand content (mass %); organic carbon content (g kg−1); bulk density (kg m−3); pH in water and depth to bedrock (cm) at 0, 5, 15, 30, 60, 100, and 200 cm depth. The first four depths, which are less than or equal to 30 cm depth, are considered as topsoil and the remaining handled as subsoil in accordance with the EU-PTFs used for calculations (Tóth et al., 2015).

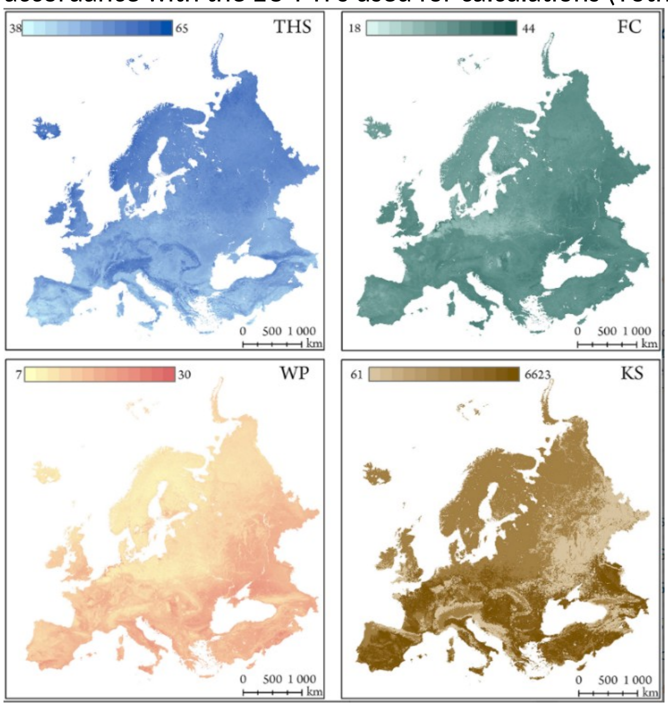


Figure 2.3.5.1. Map of FC = field capacity; KS = saturated hydraulic conductivity; THS = saturated water content; WP = wilting point.

References:

Brigitta Tóth, Melanie Weynants, László Pásztor and Tomislav Hengl, "3D soil hydraulic database of Europe at 250 m resolution", in Hydrological Processes, John Wiley & Sons Ltd, Vol.31 Issue 14, 1 July 2017, Pages 2497–2666 (pages 2662–2666); DOI: 10.1002/hyp.11203
https://esdac.jrc.ec.europa.eu/content/3d-soil-hydraulic-database-europe-1-km-and-250-m-resolution

2.4. Flux datasets for model testing

The flux datasets for model testing will be provided by the fluxnet community and ICOS. ICOS is an organisation of 12 member countries and over 130 greenhouse gases measuring stations aimed at quantifying and understanding the greenhouse gas balance of the Europe and neighbouring regions. ICOS data is openly available at the Carbon Portal, a one-stop shop for all ICOS data products.The fluxnet community summarizes all groups and study sites on the world that do eddy covariance measurements. The flux measurement sites are linked across a confederation of regional networks in North, Central and South America, Europe, Asia, Africa, and Australia, in a global network, called **FLUXNET**. This global network includes more than eight hundred active and historic flux measurement sites, dispersed across most of the world's climate space and representative biomes (**Figure 2.4.1**).

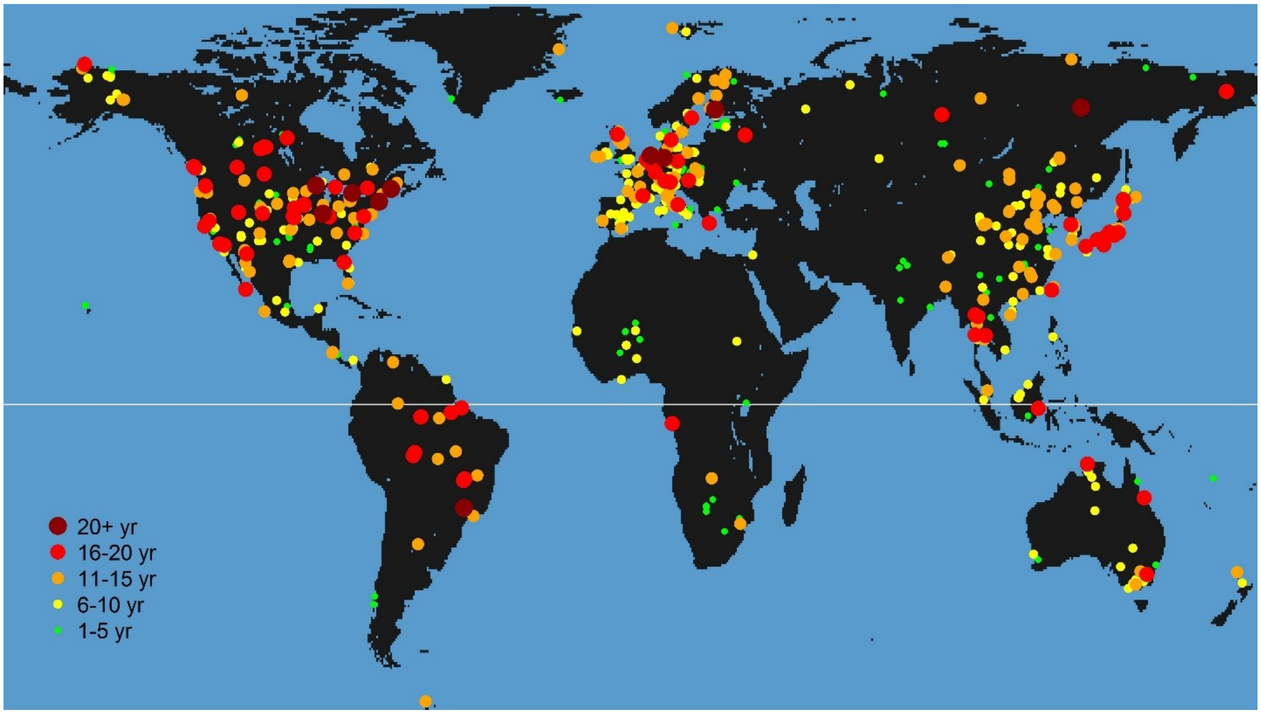


Figure 2.4.1: Measurement sites of the FLUXNET network. The different colours represent the duration of measurements.

Within this WP we detected the data demand by the modeller groups. As the models have different requirements, the demand varies e.g. for the time steps (half-hourly to daily). However, the detected target variables are:

- Gross primary production (GPP)
- Net ecosystem production (NEP)
- Respiration (TER)
- Soil moisture (SM)
- Leaf area index (LAI)
- Top of the canopy albedo

- Soil (floor) albedo
- Four-way (incoming and outgoing shortwave and longwave ?) radiation
- Photosynthetically active incoming radiation
- Sensible heat flux
- Latent heat flux
- Ground heat storage
- Transpiration
- Evaporation
- Meteorological data

Beside this list of variables meta-infromation, ancillary data (coordinates, PFT, as much information as possible describing the sites) and information on quality control (ustar filtering, other) is required. Additionally, there are often additional measurements or experiments done on the different test sites (e.g. N flux measurements, drought experiments), which would help the modellers with their work. Any of this information is detected as helpful (if available).

These data will be applied for calibration and validation of the models. Beside the data provision the teams involved in the FLUXNET community, especially the Max-Planck Institute for Biogeochemistry in Jena (Germany) will provide support with data correction and gap filling of the data sets.

2.5. Cropland management data

2.5.1. CAPRI fertilizer application data

Fertilizer application data are calculated within the CAPRI modelling framework (Britz and Witzke, 2014) using the disaggregation module (Lamboni et al., 2016; Leip et al., 2008, 2011b). For VERIFY, the disaggregation module was updated and improved.
Input data include regional statistical information from the CAPRI data base, containing regional agricultural data at the NUTS2 level for European countries, including crop areas [1000 ha], yield [kg/ha], fertilizer application rates [kg N/ha] for mineral fertilizer, applied manure and manure deposited during grazing. Fertilizer application and manure excretion rates are estimated in the CAPRI fertilizer and feed modules (Leip et al., 2011a). The data base contains also the accounting of greenhouse gas emissions (Leip et al., 2010; Weiss and Leip, 2012) and losses of reactive nitrogen (Leip et al., 2015, 2014).
The CAPDIS module provides a successive disaggregation of agricultural information from NUTS2 regional level to high-resolution spatial units (so-called HSU, homogeneous spatial units), starting with crop shares, yield and irrigation shares, livestock density, nitrogen budgets, and finally environmental indicators.

For VERIFY, the approach was improved by

- Improving a Land Area Prediction model (LAPM) to derive land use priors, based on the methods of Lamboni et al. (2016) and assessment of its accuracy (Leip et al., 2017)
- Including newly available data of crop areas from the Farm Structure Survey 2010 at a 10 km x 10 km grid (FSS2010-10k, EUROSTAT, 2017, pers. comm.), developing a methodology to gap-fill missing data which was filtered out by confidentiality rules
- Comparing crop area disaggregated to HSU level with high-resolution LPIS data for a number of countries (France, Belgium, Netherlands, Denmark) based on LAPM priors, FSS2010-10k, and only based on FSS2010-10k. The results showed that both approaches are well suited to estimate crop shares at high resolution, however, the additional information from LAPM does not significantly improve the accuracy of the predictions.
- Implementation of new data on potential and water-limited yields and update of the distribution module for yield and irrigation shares
- Update of the nitrogen distribution model

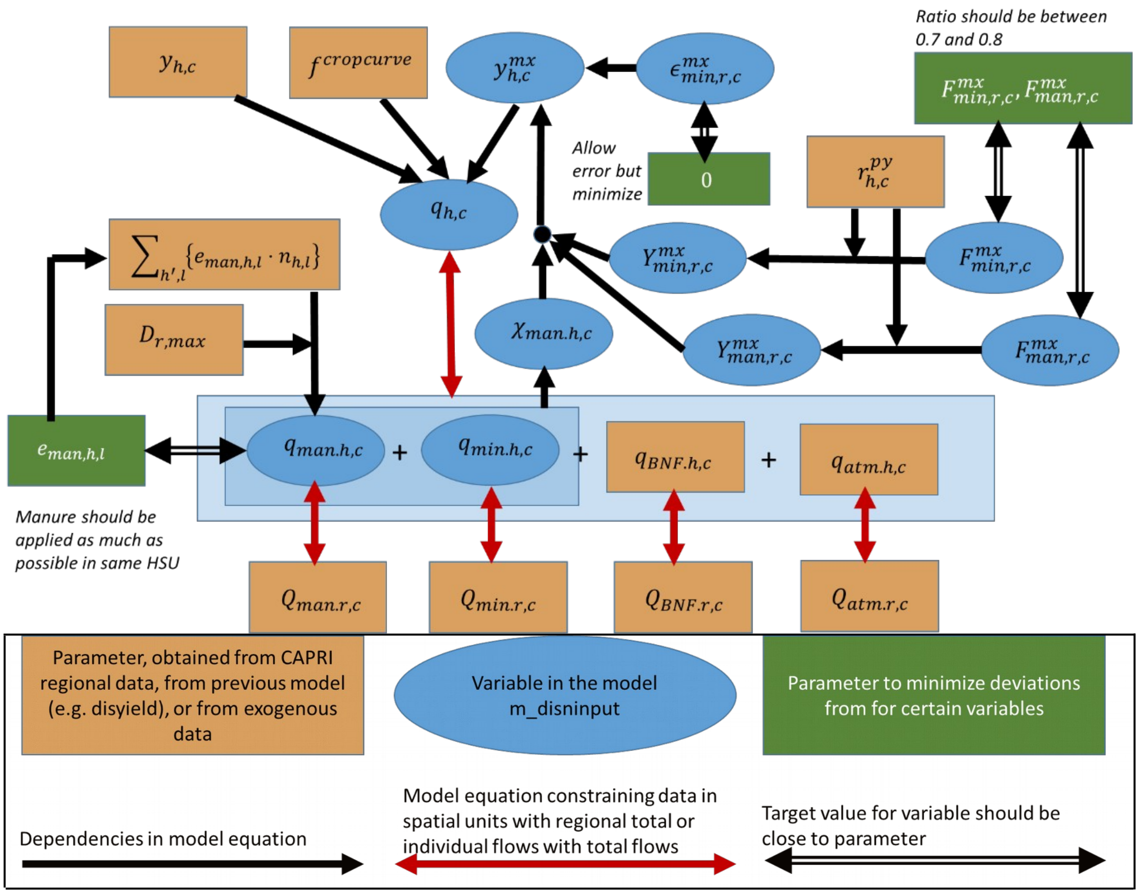


Figure 2.5.1.1. Schematic overview of the Nitrogen disaggregation model

$y_{h,c}^{\square}$ Yield [variable, kg/ha] of crop c in spatial unit h under irrigation share $f_{h,c}^{irri}$

$y_{h,c}^{mx}$ Maximum yield [variable, kg/ha] determining the shape of the crop growth curve in each $Y_{r,c}$ Yield [parameter, kg N ha^{-1} yr^{-1}] for crop c in region r.

$Y^{mx}_{r,c}$ Maximum yield [parameter, kg N ha^{-1} yr^{-1}] according to the crop response curve (Godard et al., 2008) for crop c in region r.

$Y^{mn}_{r,c}$ Minimum yield [parameter, kg N ha^{-1} yr^{-1}] according to the crop response curve (Godard et al., 2008) for crop c in region r. This parameter is set to zero in our model.

$f^{cropcurve}$ Scaling factor [parameter, dimensionless] used in the crop response curve (Godard et al., 2008). We use a uniform value of $f^{cropcurve} = 0.008$.

$Q_{r,c}$ Total N input [parameter, kg N ha^{-1} yr^{-1}] for region r and crop c.

$e_{man,h,l}$ Manure excretion [parameter, kg N/head] by animal species l in spatial unit h– net of losses in livestock housing and manure storage and management systems. No heterogeneity is assumed for nitrogen excretion rate within one NUTS2 region.

$F^{ymx}_{r,c}$ Scaling factor [variable, kg/ha] adjusting the relative potential yield so that it gives the maximum yield in the crop growth curve for each spatial unit h and crop c.

$r^{py}_{h,c}$ Relative potential yield [parameter, dimensionless] of crop c in spatial unit h.

Nitrogen input to fields needs to be distributed over the spatial units in a region where the crop is cultivated. For each crop, the regional nitrogen balance is given, with total N inputs per N input type, an N outputs or losses per output and loss type.

We base the distribution model on a generic crop-fertilizer-response curve. Such curves have the characteristics that are desirable for the disaggregation of nitrogen input:

a. higher N input leads to higher yields;
b. with increasing levels of N input the yield increment decreases;
c. the N ‘uptake’ is always less of N input.
d. saturation, attaining about 80-90% of the maximum yield at about 100-200 kg N ha^{-1} yr^{-1}

Such crop response curves assumemaximum yields, which are approached at high levels of N input. These maximum yield values are unknown. Figure 2.5.1. gives an overview of the approach indicating parameters, variables, and optimization rules.

References:

Britz, W., Witzke, P., 2014. CAPRI model documentation 2014 277.

Lamboni, M., Koeble, R., Leip, A., 2016. Multi-scale land-use disaggregation modelling: Concept and application to EU countries. Environ. Model. Softw. 82, 183–217. https://doi.org/10.1016/j.envsoft.2016.04.028

Leip, A., Billen, G., Garnier, J., Grizzetti, B., Lassaletta, L., Reis, S., Simpson, D., Sutton, M. a, de Vries, W., Weiss, F., Westhoek, H., 2015. Impacts of European livestock production: nitrogen, sulphur, phosphorus and greenhouse gas emissions, land-use, water eutrophication and biodiversity. Environ. Res. Lett. 10, 115004. https://doi.org/10.1088/1748-9326/10/11/115004

Leip, A., Britz, W., Weiss, F., De Vries, W., 2011a. Farm, land, and soil nitrogen budgets for agriculture in Europe calculated with CAPRI. Environ. Pollut. 159, 3243–3253. https://doi.org/10.1016/j.envpol.2011.01.040

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

Leip, A., Koeble, R., Rotllan-Puig, X., 2017. Land Area Prediction model (LAPM) Accuracy Assessment. European Commission, Ispra (VA), Italy. https://doi.org/10.2760/624482

Leip, A., Marchi, G., Koeble, R., Kempen, M., Britz, W., Li, C., 2008. Linking an economic model for European agriculture with a mechanistic model to estimate nitrogen and carbon losses from arable soils in Europe. Biogeosciences 5, 73–94. https://doi.org/10.5194/bg-5-73-2008

Leip, A., Wattenbach, M., Reuter, H.I., Koeble, R., Balkovic, J., Skalsky, R., Obersteiner, M., 2011b. A new data infrastructure for European terrestrial ecosystem modelling - Development of Unified Spatial Characterisation Identifier for Europe (uscie).

Leip, A., Weiss, F., Lesschen, J.P., Westhoek, H., 2014. The nitrogen footprint of food products in the European Union. J. Agric. Sci. 152, 20–33. https://doi.org/10.1017/S0021859613000786

Leip, A., Weiss, F., Wassenaar, T., Perez-Dominguez, I., Fellmann, T., Loudjani, P., Tubiello, F., Grandgirard, D., Monni, S., Biala, K., 2010. Evaluation of the livestock sector's contribution to the EU greenhouse gas emissions (GGELS) -final report. European Commission, Joint Research Centre.

Weiss, F., Leip, A., 2012. Greenhouse gas emissions from the EU livestock sector: A life cycle assessment carried out with the CAPRI model. Agric. Ecosyst. Environ. 149, 124–134. https://doi.org/10.1016/j.agee.2011.12.015

2.5.2 Management timing

The management includes beside the amounts of fertilizer and the applied practice also the dates (sowing, harvest, tillage, fertilizer and manure application). For the different techniques (e.g. tillage or fertilizer application) standard approaches are assumed, which are also often implemented in the models. The application dates will base on general assumptions and on model results.
The dates for tillage and fertilizer application will be arranged around the vegetation period (starts with sowing and ends with harvest). These dates will be simulated by phenological models, which could be implemented in the models or external. For Verify the University of Aberdeen will provide data simulated by the phenological models developed by Waha et al. (2012) and van Bussel et al. (2015). These two models are only able to provide data for wheat (spring and winter wheat) and maize. However, a literature research will allow to extend the model to more crops. At this time of the project the model is applied on previous climate data and provides data for sowing and harvest for spring wheat and maize in Europe. The next step will be to simulate by using the climate data that are provided in VERIFY (see 2.1). Additionally, the list of considered crops will be extended by using parameters provided by literature. The data are not yet uploaded on the share point platform, as the VERIFY climate data were not yet used for the phenological models.

References:
van Bussel, L. G. J., Stehfest, E., Siebert, S., Müller, C., & Ewert, F. (2015). Simulation of the phenological development of wheat and maize at the global scale. Global Ecology and Biogeography, 24(9), 1018–1029. *https://doi.org/10.1111/geb.12351*

Waha, K., van Bussel, L. G. J., Müller, C., & Bondeau, A. (2012). Climate-driven simulation of global crop sowing dates. Global Ecology and Biogeography, 21(2), 247–259. https://doi.org/10.1111/j.1466-8238.2011.00678.x

2.6. Grassland management data

The reconstructed historical maps on grassland management intensity at a resolution of 0.5° by 0.5° (1860–2012) are the minimum area of managed grassland with the fraction that is mown or grazed, yearly maps of domestic ruminant stocking density, and nitrogen (N) application rates from mineral fertilizers and manure. The minimum area of managed grassland is obtained through assuming that the grass-biomass demand equals to supply from the mown and grazed grassland at each grid cell for each year. The grass-biomass demand is derived from a livestock dataset for 2000, extended to cover the period 1860–2012. The grass-biomass supply (i.e. forage grass from mown grassland and biomass grazed) is simulated by the process-based model ORCHIDEE-GM 3.2 driven by historical climate change, rising CO_2 concentration, and changes in nitrogen fertilization.
We followed the same methods as in (Chang *et al.*, 2016) to reconstruct the history of grassland management intensity for the period 1860–2012 combining gridded and regional livestock production information with productivity from ORCHIDEE-GM v3.2. In this project, we extended the reconstructed historical grassland management intensity back to 1860 instead of 1901. To do so, we backcasted management intensity to 1860 following changes in the regional population, assuming that domestic livestock production experienced the same rate of change as population during the period 1860–1900. Another difference comes from the historical land-cover change maps. Here, the ESA CCI Land Cover product (Bontemps *et al.*, 2013) for the year 2010 was used to produce the Plant Functional Type (PFT) map used in ORCHIDEE model, following the methodology presented by (Poulter *et al.*, 2011; Poulter *et al.*, 2015). An updated release of the historical land-use forcing data set LUHv2h (http://luh.umd.edu/data.shtml; updated from LUHv1 (Hurtt *et al.*, 2011) were applied to this reference PFT map to constrain the land-cover changes of forest, grassland (combining pasture and natural grassland), and cropland during the period 1860–2010 using the backward method (BM3) following (Peng *et al.*, 2017).

References:

Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., Kirches, G., 2013. Consistent global land cover maps for climate modelling communities: current achievements of the ESA's land cover CCI. Proceedings of the ESA Living Planet Symposium, pp. 9-13.

Chang, J., Ciais, P., Herrero, M., Havlik, P., Campioli, M., Zhang, X., Bai, Y., Viovy, N., Joiner, J., Wang, X., 2016. Combining livestock production information in a process-based vegetation model to reconstruct the history of grassland management. Biogeosciences 13, 3757-3776.

Hurtt, G., Chini, L.P., Frolking, S., Betts, R., Feddema, J., Fischer, G., Fisk, J., Hibbard, K., Houghton, R., Janetos, A., 2011. Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Climatic Change 109, 117-161.

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

Peng, S., Ciais, P., Maignan, F., Li, W., Chang, J., Wang, T., Yue, C., 2017. Sensitivity of land use change emission estimates to historical land use and land cover mapping. Global Biogeochemical Cycles 31, 626-643.

Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., Zimmermann, N., 2011. Plant functional type mapping for earth system models. Geoscientific Model Development 4, 993-1010.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., 2015. Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative. Geoscientific Model Development 8, 2315-2328.

2.7. Forest management data

2.7.1 Observed diameter increment data for validation

The diameter increment data are based on freely available data from the NFIs from Germany, the Netherlands, France and Spain. For France, diameter increment data are based on increment cores while for the other countries based on repeated measurements of the same tree on permanent sample plots. See Table 2.7.1.1 for an overview of the features of the respective NFIs.

Table 2.7.1.1. Features of the NFI data set used (from Schelhaas et al. 2018a)

Country/ Region	Inventory cycle	Inventory dates	Mean measurement interval (years)	Number of plots	Plot radius (m)	Diameter threshold (cm)	Comment	NTrees
Spain	NFI2/NFI3	1986-1995/ 1996-2008	11.2	50957	5/10/15/25	7.5/12.5/22.5/42.5		557848
Netherlands	NFI5/NFI6	2001-2005/ 2012-2013	9.5	1235	variable (5-20m)	5		18348
France	NFI5-6	2005-2012	5 (core of all trees on the plot)	50404	15	7.5		474588
Germany	NFI1/NFI2	1986-1989/ 2000-2002	14.3	10344	angle count method			137425
Germany	NFI2/NFI3	2002-2012	10.2	17604	angle count method			272034

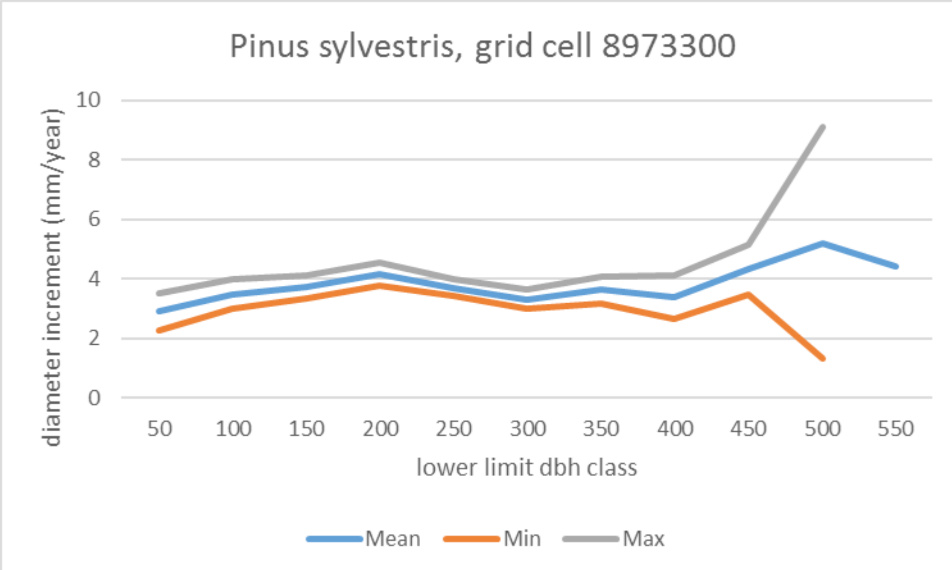


Figure 2.7.1.1. Observed mean diameter increment (mm/yr) and 95% confidence interval for Pinus sylvestris by dbh class for one grid cell.

Table 2.7.1.2 Example of validation data provided. Rastervalu corresponds to the grid cell number in the reference grid.

SimwoodSpeciesGroup	RASTERVALU	DbhClass (mm)	STD (mm)	Mean (mm)	N	ErrorMargin	Min (mm)	Max (mm)
Pinus sylvestris	8415300	200	2.001046	4.270133	75	0.452879	3.817254	4.723013
Pinus sylvestris	8416300	200	2.186867	3.613347	94	0.442094	3.171253	4.055441
Pinus sylvestris	8417300	200	1.96941	3.77855	98	0.389923	3.388627	4.168474
Pinus sylvestris	8418300	200	1.55345	3.326151	152	0.246963	3.079188	3.573114
Pinus sylvestris	8973300	200	2.454974	4.170762	162	0.378047	3.792716	4.548809
Pinus sylvestris	8974300	200	1.911555	3.744998	253	0.23555	3.509449	3.980548
Pinus sylvestris	8975300	200	2.057365	4.08785	511	0.178384	3.909466	4.266235
Pinus sylvestris	8976300	200	1.828872	4.16685	756	0.13037	4.03648	4.29722
Pinus sylvestris	8977300	200	1.71528	3.513346	970	0.107946	3.4054	3.621291
Pinus sylvestris	8978300	200	1.726381	3.711456	831	0.11738	3.594077	3.828836

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

Pinus sylvestris	8979300	200	1.950825	3.890925	294	0.222998	3.667927	4.113922
Pinus sylvestris	9532300	200	1.976968	2.450597	60	0.500242	1.950355	2.950839
Pinus sylvestris	9533300	200	2.023958	3.280083	667	0.153601	3.126482	3.433685

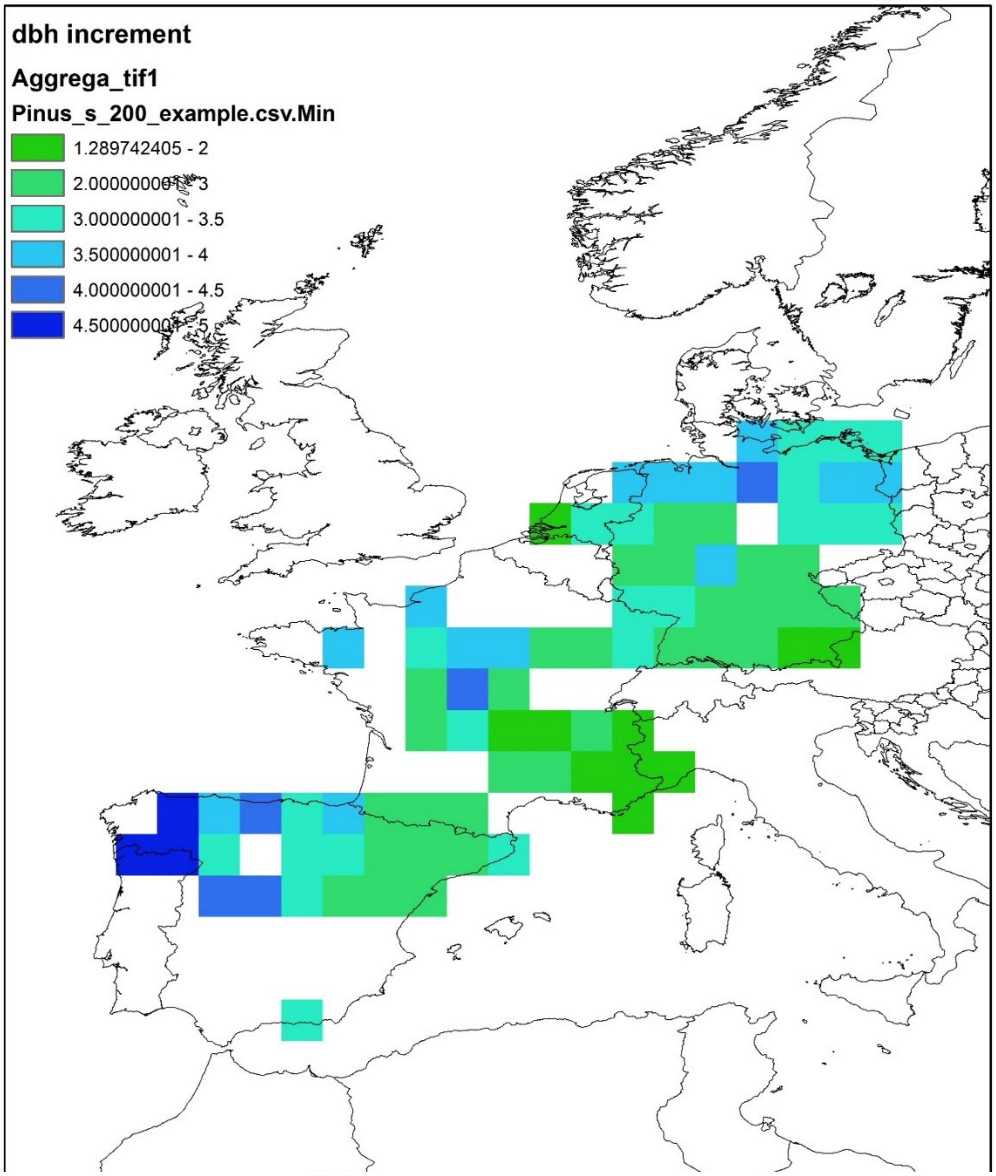


Figure 2.7.1.2 Observed mean diameter increment (mm) for Pinus sylvestris for dbh class 20-25 cm, on cells with at least 30 observations.

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

For comparison with the ORCHIDEE results we created a grid map based on the driving meteo data for ORCHIDEE, aggregated to 10x10 grid cells (CoarseGridDef.tif). This resolution seems to provide enough trees per grid cell for a reasonable estimation of dbh increment over the dbh range. Tree species are aggregated to species or species groups as described in Schelhaas et al. (2018a). Individual observations are assigned to 5 cm DBH classes based on the observed DBH at first measurement. For each combination of grid cell, tree species (group) and DBH class the average diameter increment is calculated, the standard deviation and the number of observations it is based on. Based on these, the 95% confidence interval is calculated. Table 2.7.1.2 gives an example of the data. Figure 2.7.1.1 shows the observed diameter increment in relation to the dbh for one grid cell and figure 2.7.1.2 shows an example of the spatial pattern of diameter increment.

2.7.2 Observed management and mortality data for validation

Observed mortality due to management or natural causes is based on Schelhaas et al. (2018b). The definition of the regions is given by the shapefile RegionShapeFile (Figure 2.7.3). Please note that the species grouping is different in this set. The most important species are treated individually, while the remainder is put in rest groups. No distinction in owner groups is made. Data is presented by 5 cm classes, separately for management (HarvestProbability) and natural causes (DeadProbability) (Table 2.7.3, Figure 2.7.4). Please note that harvested trees will include trees that died for natural reasons and that were subsequently extracted.

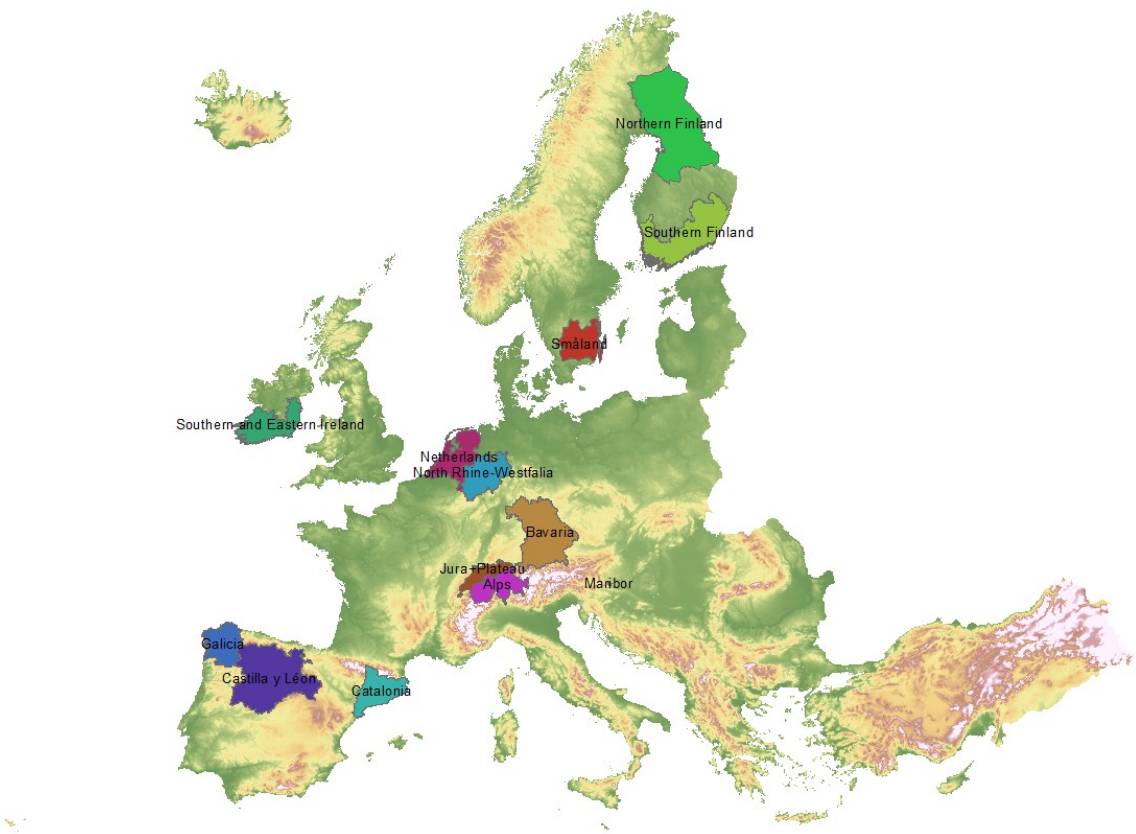


Figure 2.7.2.1. Regions for which management and mortality data is provided (taken from Schelhaas et al. 2018b).

Table 2.7.2.1. Example of mortality data for birch in the region Smaland (south Sweden)

Region	Groups	Ownergroup	DiameterClass	HarvestProbability	DeadProbability
Smaland	Betula spp.	all	0	0.037901	0
Smaland	Betula spp.	all	50	0.046305	0.000335
Smaland	Betula spp.	all	100	0.025772	0.017594
Smaland	Betula spp.	all	150	0.02753	0.017879
Smaland	Betula spp.	all	200	0.027831	0.020294
Smaland	Betula spp.	all	250	0.019405	0.022307
Smaland	Betula spp.	all	300	0.021462	0.021462
Smaland	Betula spp.	all	350	0.009051	0.013704
Smaland	Betula spp.	all	400	0.009261	0.009261

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

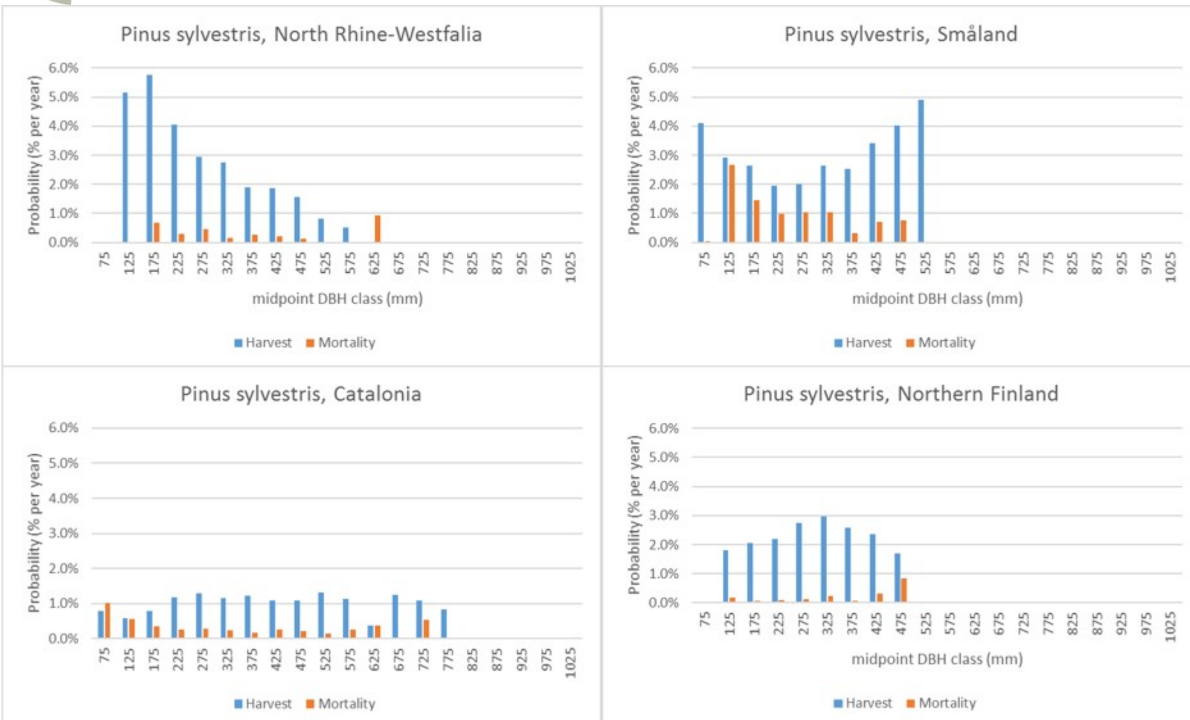


Figure 2.7.2.2. Example of regionally different patterns of management and mortality in Pinus sylvestris (Schelhaas et al. 2018b).

References:

Schelhaas MJ, Hengeveld GM, Heidema N, Thürig E, Rohner B, Vacchiano G, Vayreda J, Redmond J, Socha J, Fridman J, Tomter S, Polley H, Barreiro S, Nabuurs GJ, 2018a. Species-specific, pan-European diameter increment models based on data of 2.3 million trees. Forest Ecosystems 5:21 doi.org/10.1186/s40663-018-0133-3

Schelhaas MJ, Fridman J, Hengeveld GM, Henttonen H, Lehtonen A, Kies U, Krajnc N, Lerink B, Ní Dhubháin A, Polley H, Pugh TAM, Redmond J, Rohner B, Temperli C, Vayreda J, Nabuurs GJ, 2018b. Actual European forest management by region, tree species and owner based on 714,000 re-measured trees in national forest inventories. PLoS ONE 13(11): e0207151.

2.8. Nitrogen deposition data

Several models related to carbon cycle and N2O indirect emissions need gridded atmospheric deposition fields, mainly of oxidized and reduced nitrogen (NOy and NHx). There are several model products for nitrogen deposition available. The aim of D.1 is to provide a harmonized currently best available and fit for purpose gridded estimates of N deposition (database) to the various modelling activities (the latter including the tiered approach to estimate indirect effects on N2O).

VERIFY has as an overarching goal to provide a pre-operational GHG verification system, that after the project's end may become operational. Therefore, it is of high importance to liaise this component to international activities that aim to provide operationally high-quality deposition products. In this context, the WMO's Global Atmospheric Watch has put in place the MMF-TAD project (Model-Measurement-

Fusion, Total Atmospheric Deposition), which aims to provide high-quality regional and global deposition products, based on model-measurement fusion approaches to a variety of users. (Carou et al., 2017; Vet et al., 2014)

The short-term objectives are to provide recommended global and regional gridded datasets, quality controlled with observations, and in the long-term to include into the Copernicus CAMS model the data-assimilation and fusion approaches that are explored in MMF at regional scales. Since the MMF-TAD project, for which currently the schedule of deliverables is not yet defined, will follow a different development schedule compared to Verify in this stage we recommend to use the deposition data products that are associated with MMF-TAD and readily available.

The initial database for deposition estimates, linked to the MMF-TAD project consists of:

Global deposition centered around the year 2010:

- HTAP global ensemble (Tan et al., 2018)
- CCMI/CMIP6 2010 timeslices (M. Hegglin, manuscript in preparation, 2019).
- EMEP global model for 2010 (Schwede et al., 2018)

Regional deposition for Europe:
- EMEP regional model results (www.emep.int)

Results from the TFMM & AQMEII model inter-comparison (Vivanco et al., 2018)

Results for time slices back in time and for a variety of scenarios:
Global:
- CCMI/CMIP6 2010 time slices (M. Hegglin, manuscript in preparation) or CMIP5: (Lamarque et al., 2013)

Europe:
- EMEP regional model back to 1990.

These model results, which are part of the MMF-TAD project, will be used for initial sensitivity studies, aiming at understanding the sensitivity of impact models for variations in deposition fields. In a next step the model products will be updated according to user requirements and priorities.

An example of a global deposition gridded dataset is a downscaled high-resolution dataset, taking into account sub-grid forest fraction, using the the EMEP model (Figure 2.8.1) which was recently published by Schwede et al. (2018).

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

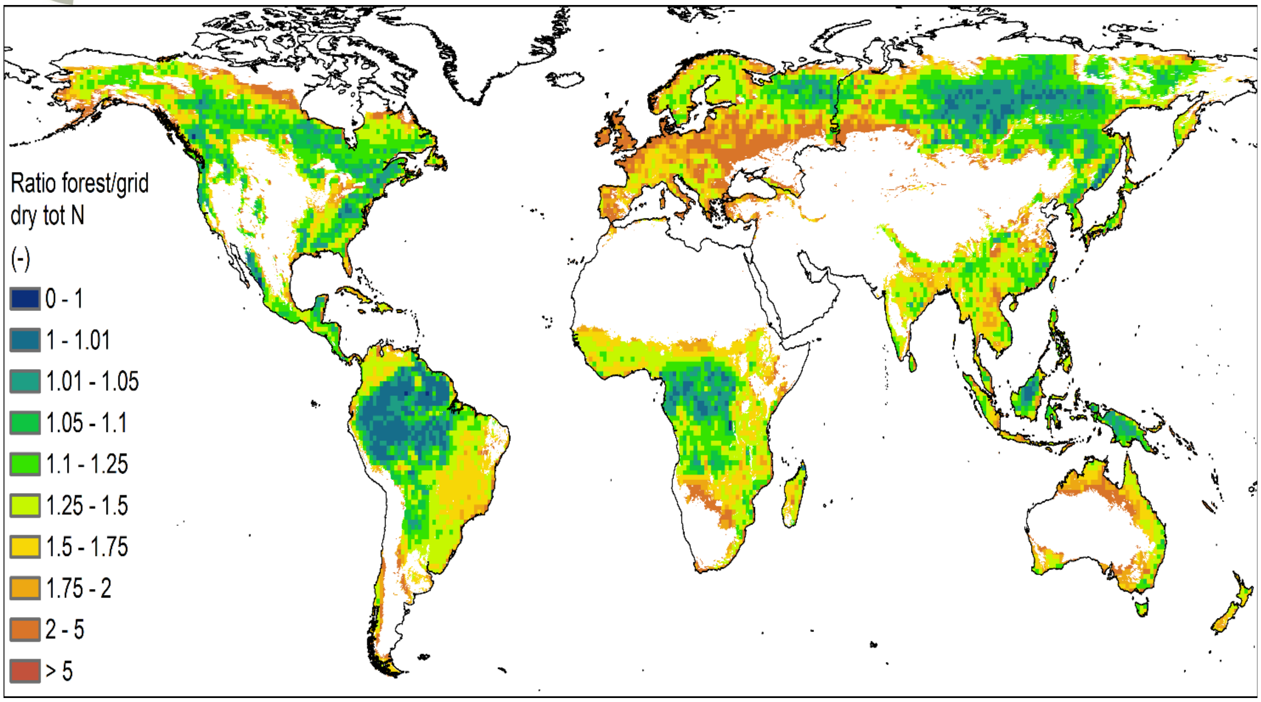


Figure 2.8.1: EMEP model (Schwede et al, 2018;). Global model that delivers ecosystem (forest) specific deposition of nitrogen deposition. In some regions relative large differences are found with HTAP2 global ensembleIn particular in mixed forest systems, due to large enhancement of dry deposition.

2.9. Freshwater fluxes and river exports

This data set represents a climatology of average annual CO_2 emissions from lakes (incl. reservoirs) and rivers at the spatial resolution of 0.1°. For lakes, we used the empirical model developed by Hastie et al. (2018, GCB) to predict the average lake CO_2 partial pressure (pCO_2) per 0.1° grid cell, which was then combined with data on lake surface area and alternative estimates of gas exchange velocity to estimate the lake-atmosphere CO_2 flux (see Hastie et al., 2018, GCB for details). The lake surface areas were taken from the HydroLAKES database (Messager et al., 2016, Nat Comm). Riverine CO_2 emission estimates were downscaled from the global dataset by Lauerwald et al. (2015, GCB) (from 0.5° to 0.1°), using high resolution estimates of river surface area based on the HydroSHEDs database (Lehner et al., 2008, Eos).
In addition, the climatology gives the 5th and 95th% percentile of uncertainty range in the CO_2 emission estimates, which were derived from Monte Carlo simulations (10,000 runs) propagating the uncertainties related to the estimated water-atmosphere pCO2 gradient, gas exchange velocity and surface water area. For the area covered by the NUTS 2016 regions (EU membership countries, + EU candidates and EFTA countries), we estimate an annual emission of 15.7 (4.7-31.3) Tg C yr^{-1} from lakes and of 14.7 (10.5-19.8) Tg C yr^{-1} from rivers.
For the next year release (D3.2), the climatology will be upgraded to additionally include seasonality in CO_2 emissions and CO_2 emission from estuaries.

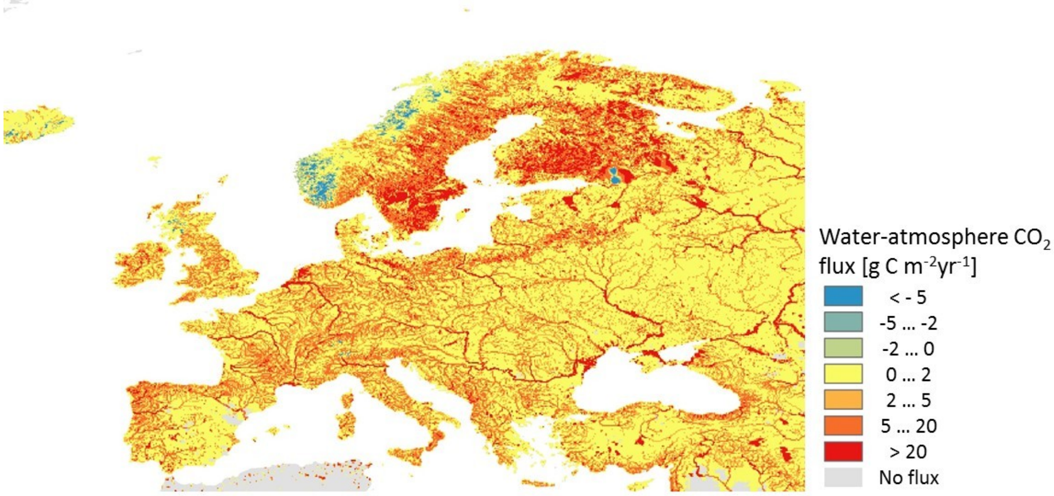


Figure 2.9.1: Estimated average annual water-atmosphere CO_2 flux. Emission rates refer to continental area (terrestrial + aquatic).

References:
Hastie, A., Lauerwald, R., Weyhenmeyer, G., Sobek, S., Verpoorter, C., & Regnier, P. (2018). CO_2 evasion from boreal lakes: Revised estimate, drivers of spatial variability, and future projections. Global Change Biology, 24(2), doi: *10.1111/gcb.13902*.

Lauerwald, R., Laruelle, G. G., Hartmann, J., Ciais, P., & Regnier, P. A. G. (2015). Spatial patterns in CO_2 evasion from the global river network. Global Biogeochemical Cycles, 29(5), 534–554, doi: 10.1002/2014GB004941.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. Eos, Transactions, AGU, 89(10), 93–94.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nature Communications, 7, 13603, doi: 10.1038/ncomms13603.

2.10. Coastal ocean CO2 fluxes

Air-sea CO_2 fluxes have been estimated for the following coastal regions: the Baltic, the North Sea, the Norwegian coastline, and the Barents Sea, defined by using the 500 m isobath. The maps cover the period 1997-2016 and are provided at a resolution of 0.125°x0.125°. The underlying sea surface pCO_2 maps were created by optimising the global product of Rödenbeck et al. (2013) in the coastal zones. The advantage is that this creates flux estimates that are consistent with those of the open ocean product, in particular with respect to time trends. They are also significantly more accurate than the original global product. Briefly, coastal pCO_2 observations were extracted from SOCAT (Bakker et al., 2016) and fitted to collocated data for chlorophyll *a*, sea surface temperature, sea-surface salinity, mixed layer depth, ice concentration, and gridded pCO_2 from Rödenbeck et al. (2013). The resulting multilinear regressions (MLRs) were next applied to the gridded data for the full regions mentioned above, to obtain mapped pCO_2 fields. The resulting fields agrees significantly better with the observations than the original

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

Rödenbeck estimates (e.g. North Sea RMSE 26 μatm vs 95 ματμ for Rödenbeck). This can be attributed to the higher spatial resolution (0.125° vs. 4°x5° for Rödenbeck) and the MLRs' better representation of specific coastal processes (freshwater effects, primary production, etc.). The MLRs that included the Rödenbeck pCO_2 values among the independent variables performed better than MLRs that didn't (North Sea RMSE for a 'free' MLR was 32 μatm), and also better than an MLR that included pCO_2 from the global Landschützer et al. (2013) product among the independent variables (North Sea RMSE 36 μatm). This is likely the result of the Landschützer et al. (2014) maps not covering coastal regions, in contrast to the Rödenbeck et al. (2013) maps. The procedures have been described in Becker et al. (in prep). This approach will now be scaled up to the entire European coastal zone.

References:

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W. J., Castle, R. D., Chen, L. Q., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C., Huss, B., Ibanhez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Kortzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschutzer, P., Lauvset, S. K., Lefevre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S. Q.: A multi-decade record of high-quality fCO_2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT), Earth Syst Sci Data, 8, 383-413, 2016.

Becker, M., Olsen, A., Landschützer, P., Omar, A., Rödenbeck, C., and Skjelvan, I.: Mapping coastal fCO_2 using multilinear regression, in preperation.

Landschützer, P., Gruber, N., Bakker, D. C. E., and Schuster, U.: Recent variability of the global ocean carbon sink, Global Biogeochem Cy, 28, 927-949, 2014.

Rödenbeck, C., Keeling, R. F., Bakker, D. C. E., Metz, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean p(CO2) and sea-air CO_2 flux variability from an observation-driven ocean mixed-layer scheme, Ocean Sci, 9, 193-216, 2013.

2.11. XCO2 from OCO-2 via FOCAL algorithm

The fast atmospheric trace gas retrieval for OCO2 (FOCAL-OCO2) has been setup to retrieve XCO2 (the column-average dry-air mole fraction of atmospheric CO2) by analyzing hyper spectral solar backscattered radiance measurements of NASA's OCO2 satellite. FOCAL includes a radiative transfer model which has been developed to approximate light scattering effects by multiple scattering at an optically thin scattering layer. This reduces the computational costs by several orders of magnitude. FOCAL's radiative transfer model is utilized to simulate the radiance in all three OCO-2 spectral bands allowing the simultaneous retrieval of CO2, H2O, and solar induced chlorophyll fluorescence.

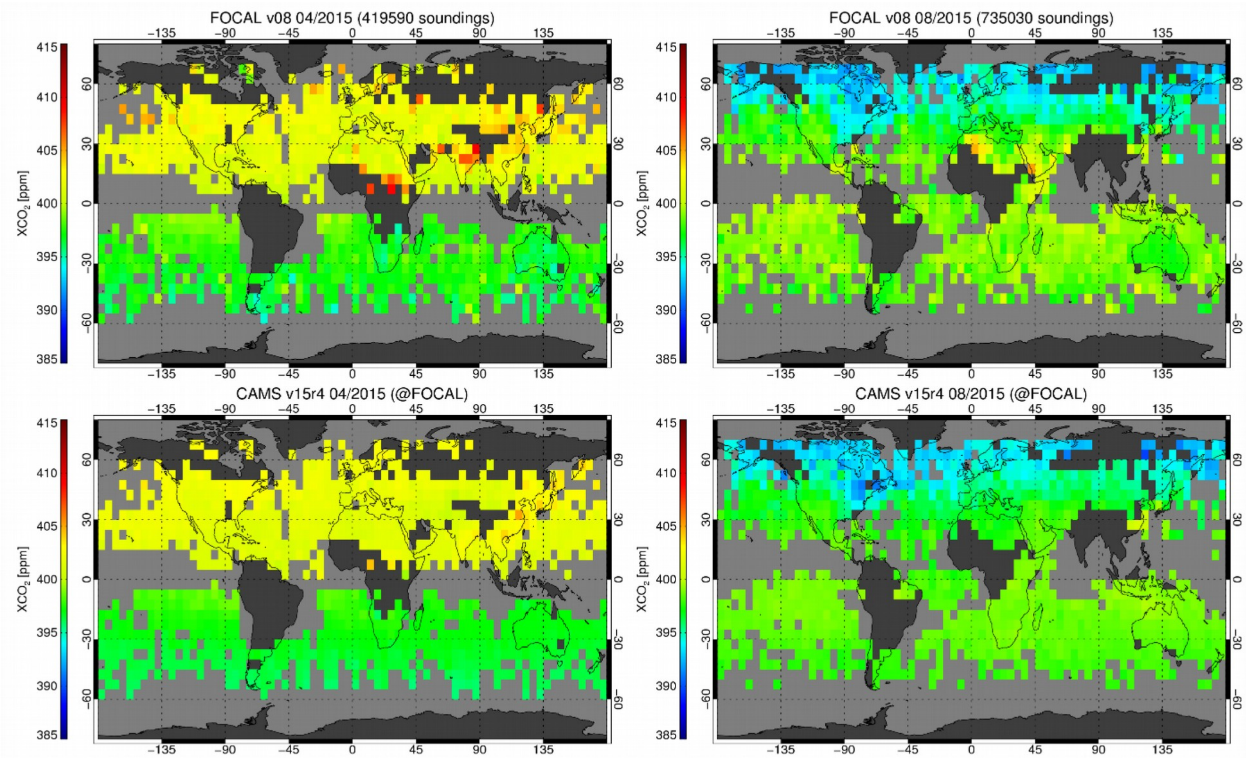


Figure 2.11.1- Monthly mean XCO2 at 5° x 5°. Top: FOCAL v08. Bottom: CAMSv15r4 sampled as FOCAL. Left: April 2015. Right: August 2015.

A first FOCAL version 08 data set has been generated and made available for VERIFY. This data set covers the time period 2015-2016.

The data product has been compared with Total Carbon Column Observation Network (TCCON) ground based XCO2 retrievals. The validation results are as follows: bias (standard deviation of site-to-site biases): 0.58 ppm, scatter (single measurement precision): 1.50 ppm.

Details see "Data Product User Guide forXCO2OCO-2/FOCAL (v08)":
http://www.iup.uni-bremen.de/~mreuter/TN_XCO2-OCO2-FOCAL_v08.pdf

References:

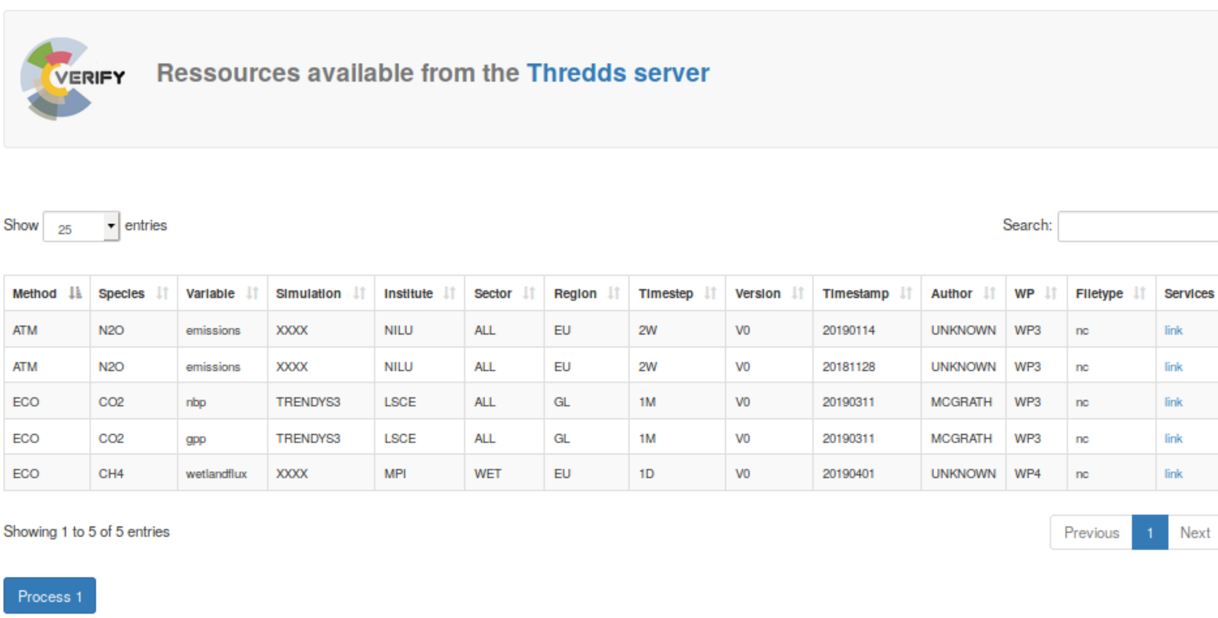
M.Reuter, M.Buchwitz, O.Schneising, S.Noel, V.Rozanov, H.Bovensmann and J.P.Burrows: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering - Part 1: Radiative Transfer and a Potential OCO-2 XCO2 Retrieval Setup Remote Sensing, 9(11), 1159; doi:10.3390/rs9111159, 2017a

M.Reuter, M.Buchwitz, O.Schneising, S.Noel, H.Bovensmann and J.P.Burrows: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering - Part 2: Application to XCO2 Retrievals from OCO-2 Remote Sensing, 9(11), 1102; doi:10.3390/rs9111102, 2017b

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

3. Organization of the database

A list of available data and the data itself will be available from the VERIFY THREDDS data server (TDS, https://verifydb.lsce.ipsl.fr/thredds/catalog.html), with some limited metadata available from the VERIFY data catalogue (available from the VERIFY web site: http://verify.lsce.ipsl.fr/index.php/products). Note that for the VERIFY partners additional information on the different data sets is available under the password protected share point platform (https://projectsworkspace.eu/sites/VERIFY/Lists/WP3inputdataset/AllItems.aspx).
The filenames will be assigned to contain various information about the file itself, including the method, species, institute, region, spatial coverage, temporal resolution, and the persion who uploaded the file. This information is used to automatically generate a catalogue of available data (see figure below). The TDS is developed by Unidata, a member of the UCAR Community Programs, managed by the University Corporation for Atmospheric Research, and funded by the National Science Foundation of the United States, with the goal of helping educators and scientists obtain and use geoscience data. The TDS also supports several dataset collection services including some sophisticated dataset aggregation capabilities. This allows the TDS to aggregate a collection of datasets into a single virtual dataset, greatly simplifying user access to that data collection. The TDS also contains viewing tools to facilite direct user browsing of stored datasets, instead of forcing the user to rely on metadata.


Ressources available from the Thredds server

Show 25 entries Search:

Method	Species	Variable	Simulation	Institute	Sector	Region	Timestep	Version	Timestamp	Author	WP	Filetype	Services
ATM	N2O	emissions	XXXX	NILU	ALL	EU	2W	V0	20190114	UNKNOWN	WP3	nc	link
ATM	N2O	emissions	XXXX	NILU	ALL	EU	2W	V0	20181128	UNKNOWN	WP3	nc	link
ECO	CO2	nbp	TRENDYS3	LSCE	ALL	GL	1M	V0	20190311	MCGRATH	WP3	nc	link
ECO	CO2	gpp	TRENDYS3	LSCE	ALL	GL	1M	V0	20190311	MCGRATH	WP3	nc	link
ECO	CH4	wetlandflux	XXXX	MPI	WET	EU	1D	V0	20190401	UNKNOWN	WP4	nc	link

Showing 1 to 5 of 5 entries

Previous 1 Next

Process 1

Figure 3.1: Illustration of the catalogue that will display the available product gathered/produced within VERIFY.

As the data come from different sources and the data are of different nature, the format of the data sets varies. However, data that are produced or aggregated within this project will be provided as netcdf-files or csv-files. This is required due to data sets being available at different spatial and temporal resolutions. As the resolution varies due to different origin, data basis and purposes of the data sets, newly produced

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.

or aggregated data will be provided in an agreed-upon resolution of 0.125 degree grid maps to the extent possible. As several modelling groups have already their own aggregation tools (partially implemented in the model approaches) a central harmonization is not yet required.

At this point of the project the data are collected, summarized and described as they are available. The challenge for the data organization is the parallel improvement and use of the different data sets (e.g. climate data of task 2.1). Additionally, certain data sets will be adapted to the demands of the modelling groups (e.g. the land use/land cover data set of task 2.2). Therefore, teleconferences were organized to coordinate the demand of data and progress of development (e.g. the teleconference regarding the progress of the land use/land cover data of task 2.2). There is an ongoing discussion about the nitrogen deposition data set of task 2.8. Most data sets are available, but the optimum format is not yet finalized (see details in section 2.8). The next description of the database (follow up deliverable) will provide a more comprehensive and fully standardized database.

VERIFY is a research project funded by the European Commission under the H2020 program. Grant Agreement number 776810.