Ancillary Data for Numerical Prediction Models at Météo-France

Patrick Le Moigne[(1)], Bertrand Decharme[(1)], Marie Minvielle[(1)], Simon Munier[(1)], Diane Tzanos[(1)], Jean Wurtz[(1)]

[(1)] CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France.

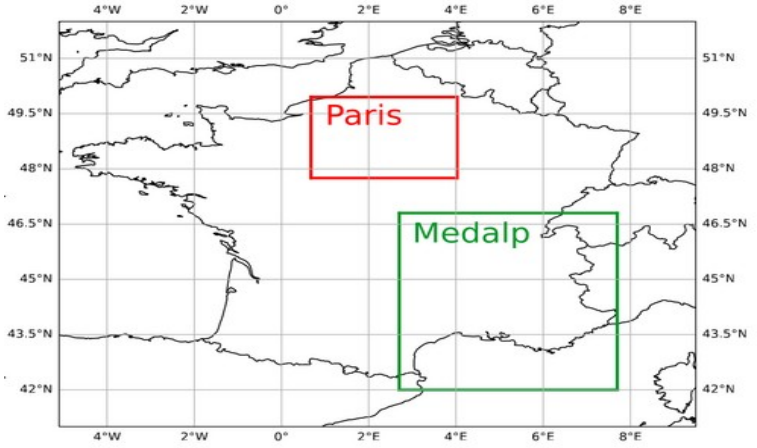
Outline

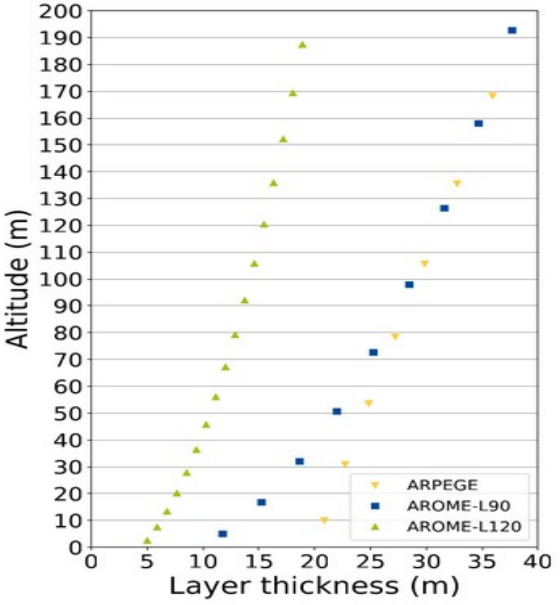
- Introduction
- Ancillary data for SURFEX
- Ongoing work

Introduction

- At Météo-France, as elsewhere, model resolution is increasing to better represent local processes, induced by the interactions due to landscape surface heterogeneities (BL turbulence, breezes, etc.). Recent work is approaching the hectometric scale for limited area models.

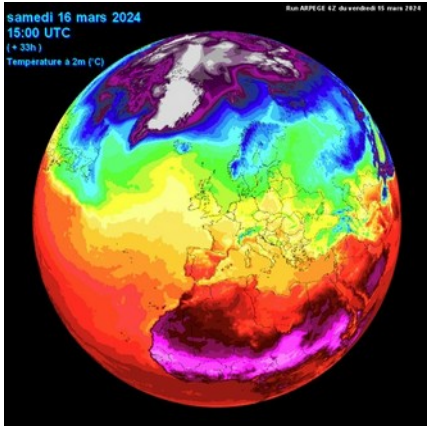
- ✔ 2 configurations Paris (36h) and Medalp (24h)
- ✔ Dynamical Adaptation from AROME-France
- ✔ Operational since Autumn 2024

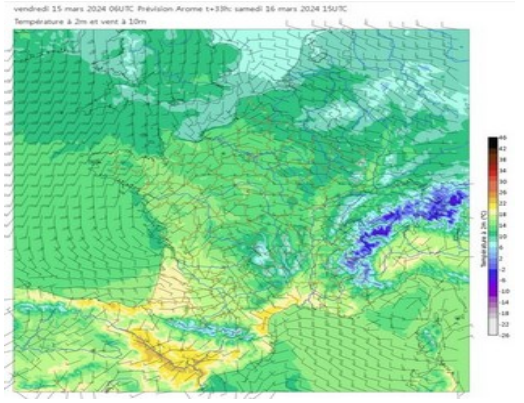




Introduction

- Météo-France's operational models are the global ARPEGE model (stretched grid, 7.5km over France) and the LAM model AROME (1.3km over France). Climate applications rely on the global CNRM-CM model. High resolution research models are AROME and Méso-NH (LES).

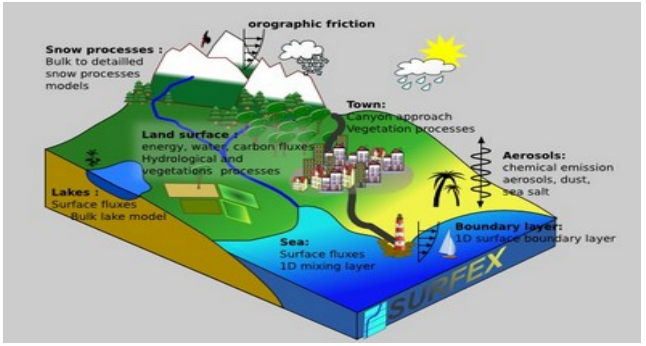


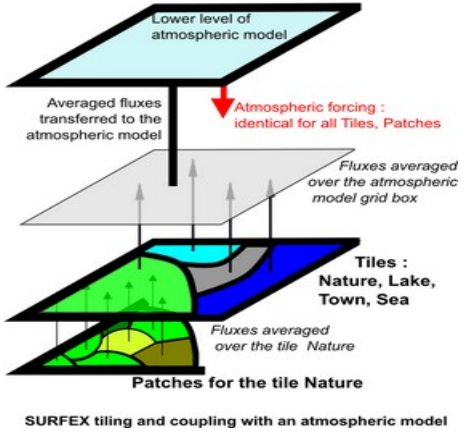


- Mandatory to accurately define model parameters (ancillary data) at high spatial resolution and have the possibility to update them in time to improve the realism of our models.

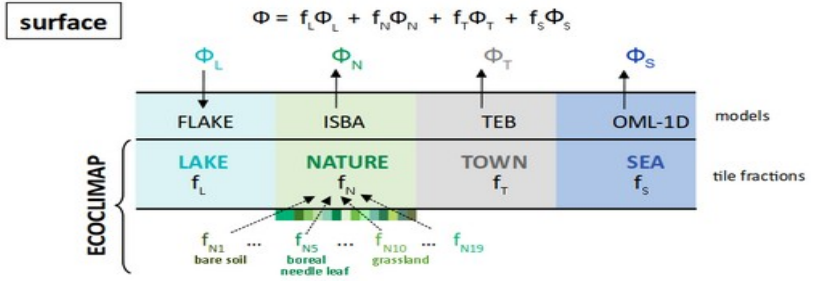
The SURFEX land surface modelling platform

- Surface modelling is based on SURFEX for all NWP and climate models at MF.
- Tiles (and subtiles for nature) used to capture the landscape heterogeneity.
  https://www.umr-cnrm.fr/surfex/





SURFEX tiling and coupling with an atmospheric model



Ancillary Data

- **Model parameters**
- **Tiles fractions**
- **Topography**
- **Soil texture**
- **...**

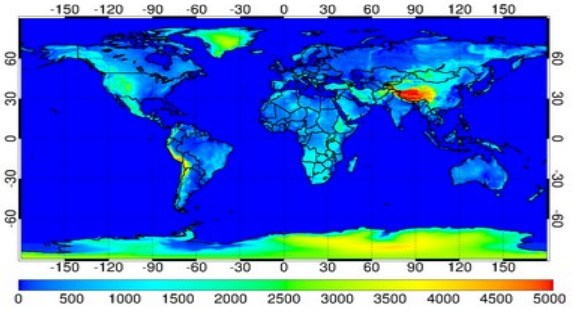
Ancillary data for SURFEX
What do we need? What for?

Topography, Depth, Bathymetry

Oceans/Seas

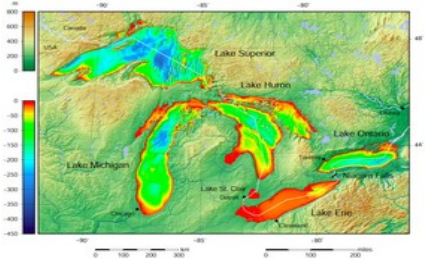
- Bathymetry **ETOPO2** (NOAA)
- CNRM-CM model coupled to NEMO ocean model
- 1D Ocean Mixed Layer model

Continental surfaces

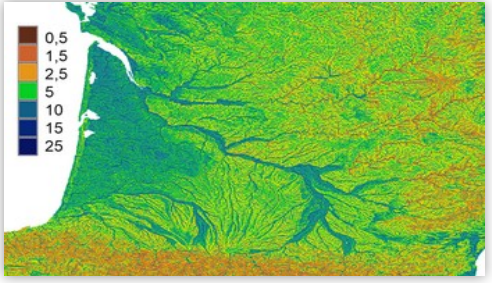


- Topography **GMTED2010** (USGS&NGA)
- Vertical interpolation (climatological gradient)
- Subgrid scale orography (slope, anisotropy)
  - Roughness length
  - Envelope

Lakes



- Depth **GLDBv1**
- Light extinction coefficient
- Bathymetry when available
- FLake model

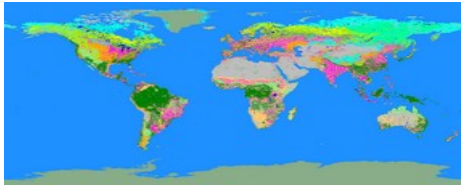


- Topographic index for runoff (HydroSHEDS)

Ancillary data for SURFEX
What do we need? What for?

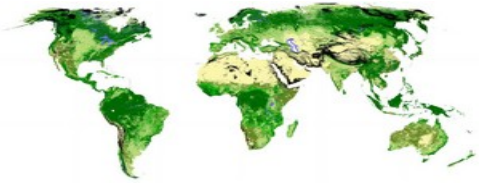
Land cover description: ECOCLIMAP

- CNRM home-made global database of land covers
- It describes the types of surface covering the whole Earth, and associated land surface parameters
- ECOCLIMAP-I 1km resolution with 215 covers (combination of vegetation types), ECOCLIMAP-II +273 covers for Europe
- ECOCLIMAP-SG 300m resolution, based on ESA-CCI LC product (pure cover)

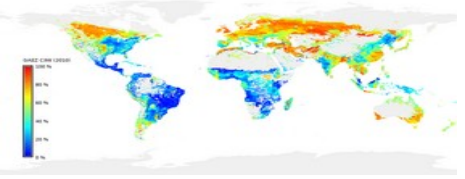


Land cover ECOCLIMAP-SG

- LAI (CGLS 2014-2016)
- Albedo (CGLS 300m → 2020, MODIS 500m)
- Height of trees (Global Canopy Height, ETH Zürich)



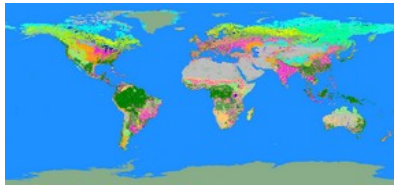
LCZ generator + GeoClimate/OSM



Winter C3 crops (Global Agro-Ecological Zones)

Ancillary data for SURFEX
What do we need? What for?

Land cover description: ECOCLIMAP



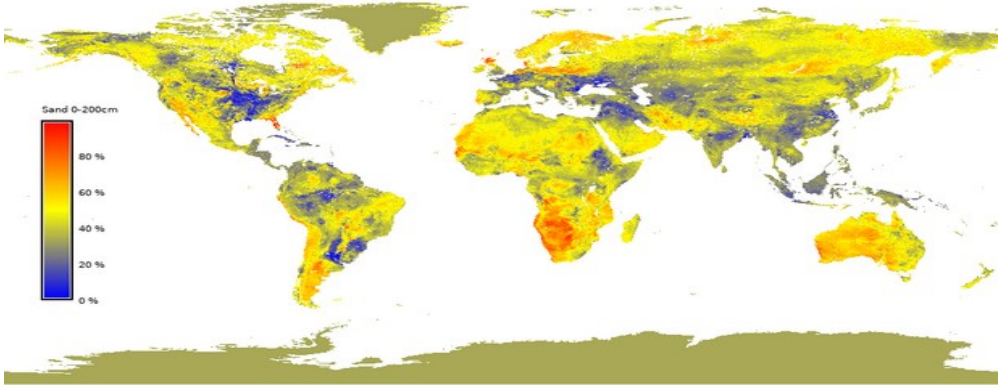
- To define tiles, sub-tiles
- To attribute parameters by tiles/sub-tiles, time

33 covers of ECOCLIMAP-SG					
1	Sea and oceans	**12**	Boreal needleleaf evergreen	**23**	Flooded grassland
2	lakes	**13**	Temperate needleleaf evergreen	**24**	LCZ1 : compact high-rise
3	rivers	**14**	Boreal needleleaf deciduous	**25**	LCZ2 : compact midrise
4	Bare land	**15**	shrubs	**26**	LCZ3 : compact low-rise
5	Bare rock	**16**	Boreal grassland	**27**	LCZ4 : open high-rise
6	Permanent snow	**17**	Temperate grassland	**28**	LCZ5 : open midrise
7	Boreal broadleaf deciduous	**18**	Tropical grassland	**29**	LCZ6 : open low-rise
8	Temperate broadleaf deciduous	**19**	Winter C3 crops	**30**	LCZ7 : lightweight low-rise
9	Tropical broadleaf deciduous	**20**	Summer C3 crops	**31**	LCZ8 : large low-rise
10	Temperate broadleaf evergreen	**21**	C4 crops	**32**	LCZ9 : sparsely built
11	Tropical broadleaf evergreen	**22**	Flooded trees	**33**	LCZ10 : heavy industry

Ancillary data for SURFEX
What do we need? What for?

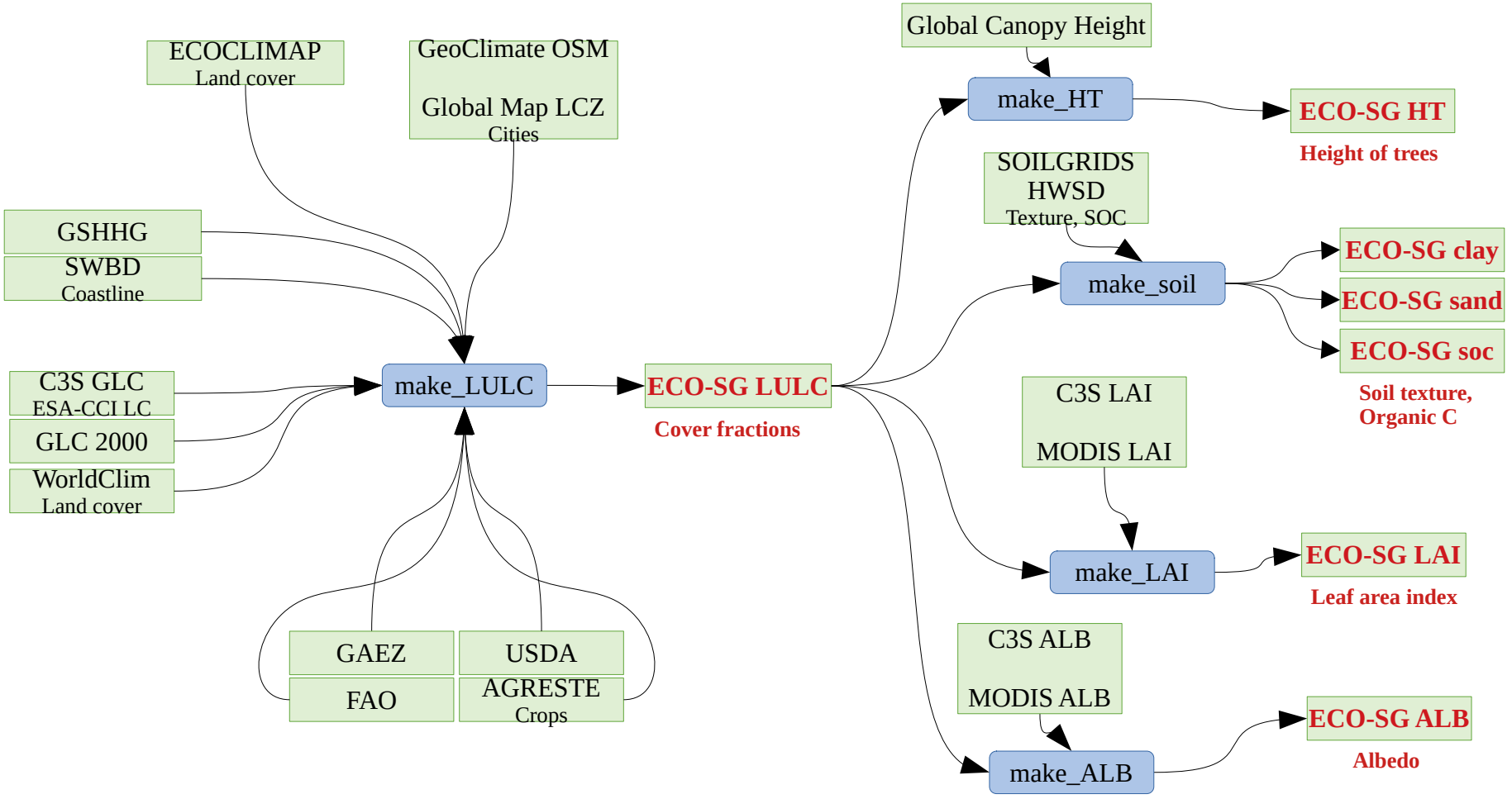
Sub-surface parameters

- Soil texture (clay, sand fractions), HWSD, SOILGRIDS
  - Used to calculate specific parameter of the **ISBA** land surface model
  - Soil moisture at field capacity, wilting point, saturation
  - Hydraulic conductivity at saturation
  - ...
- Soil Organic Carbon



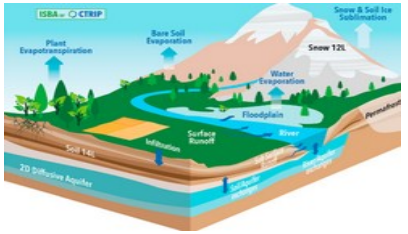
Sand fraction from SOILGRIDS dataset

ECOCLIMAP Algorithm

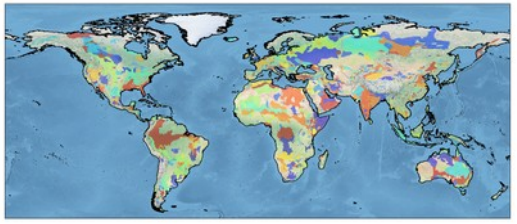


Ancillary data for ISBA-CTRIP coupled system

Hydrology: ISBA-CTRIP model



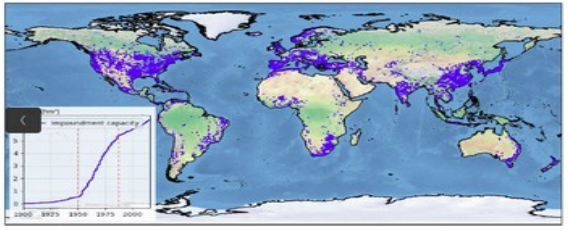
- River network (MERIT-Hydro)
- Sub-grid topography (aquifers, floodplains, MERIT-DEM)
- River morphology (length, slope, roughness)



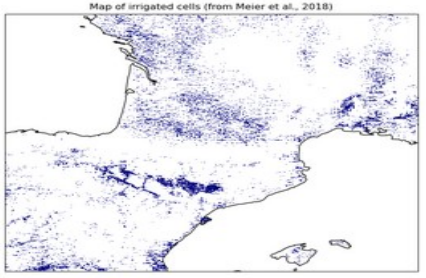
Aquifers (WHYMAP)

River network 1/12° (MERIT-Hydro)

- Anthropization

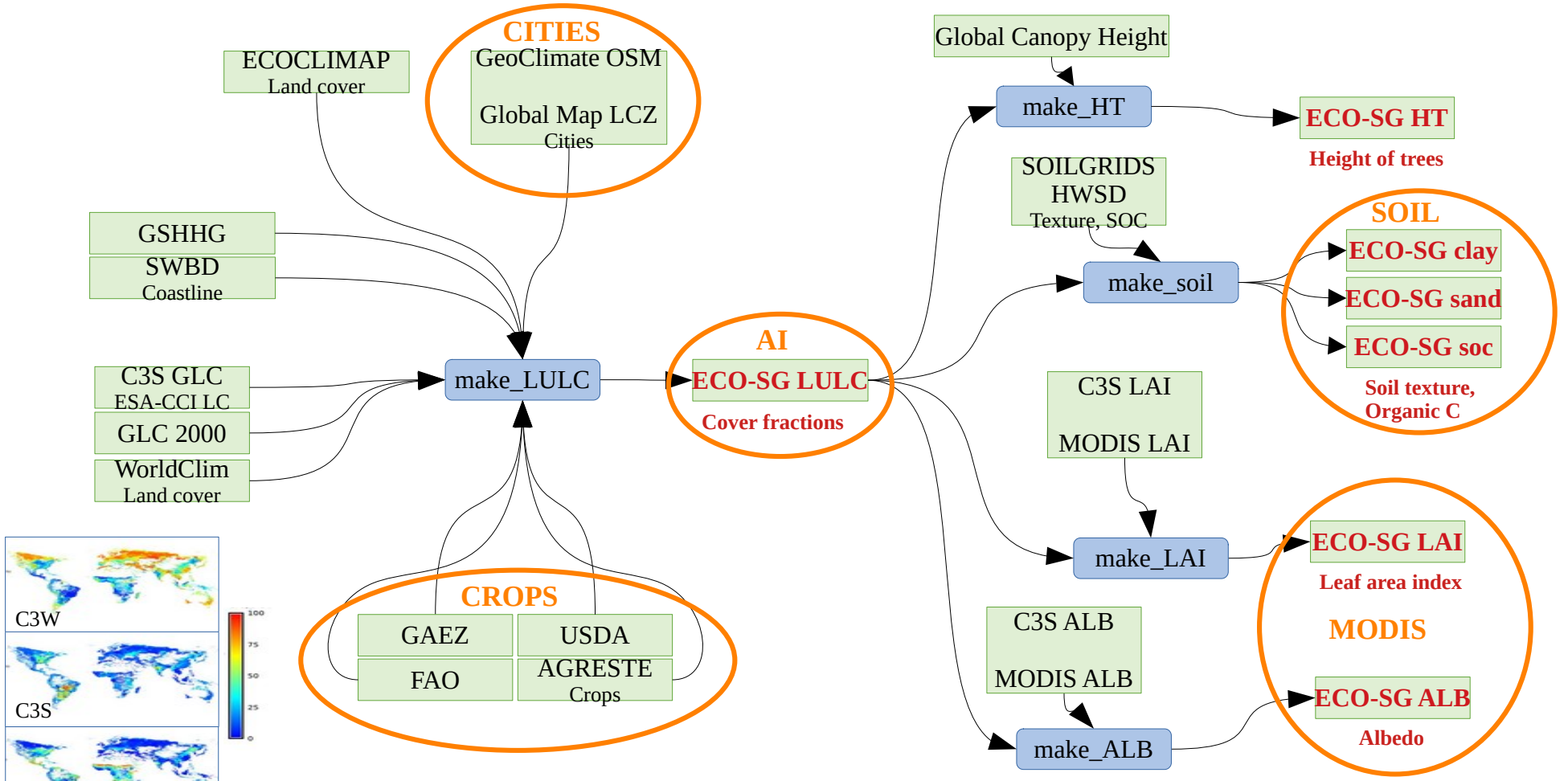


Dams-reservoirs distribution (GRanD)



Irrigated cells (south France)

Ongoing work



Ongoing work: CITIES

- **GeoClimate** tool to create Urban Canopy Parameter maps

GeoClimate preparation algorithms

GeoClimate geoindicators algorithms

Vector data

- French National Topographic data
- OpenStreetMap

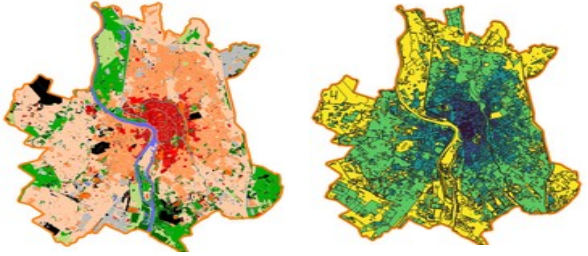
GeoClimate standard inputs

Buildings
Vegetation
Impervious
Water
Road
Rail
Zone
Sea-land
Urban areas

GeoClimate standard outputs

LCZ classification

Others UCP

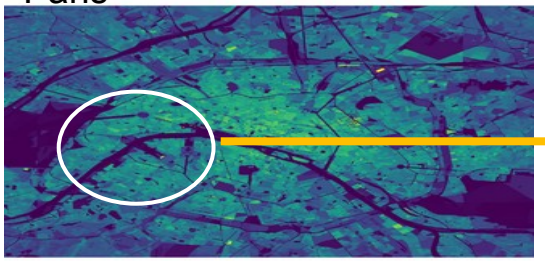


Building SVF
Facade density
Land fraction
Building types...

Ongoing work: CITIES

- **GeoClimate / OSM**: a high resolution database

Paris



Le Trocadero

- An **AI** method used to estimate building height from **OSM** data when missing

« Independent variables » are calculated (building, block and RSU scales – size, form and land-type fractions)

« Dependent variable » from a reference dataset is used as TRUE building height

RandomForest model to estimate building height

Buildings with estimated height (m)

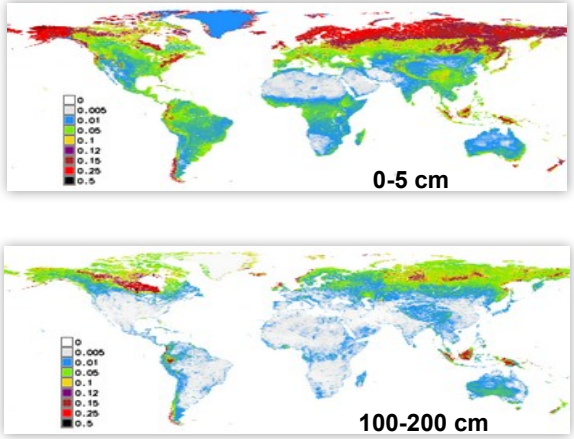
Method evaluated only in France !

Ongoing work: revision of soil physical properties for ISBA

- **Soil : SoilGrid2.0** at 250m with 6 horizons (0-5, 5-15, 15-30, 30-60, 60-100, 100-200 cm)
  - ✔ Sand & Clay (kg.kg^{-1})
  - ✔ Soil bulk density (kg.m^{-3})
  - ✔ Organic carbon content (kg.kg^{-1})
  - ✔ Coarse fragments (m^3.m^{-3})

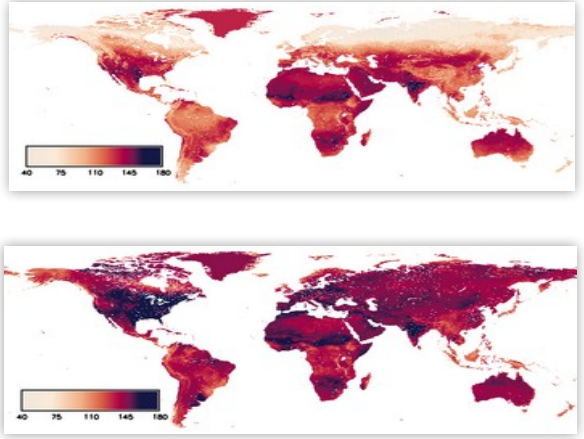
Interpolations on ISBA soil grid with boundaries (up and down) extrapolations accounting for the shape of the actual profile

Soil Organic carbon content (kg.kg^{-1})



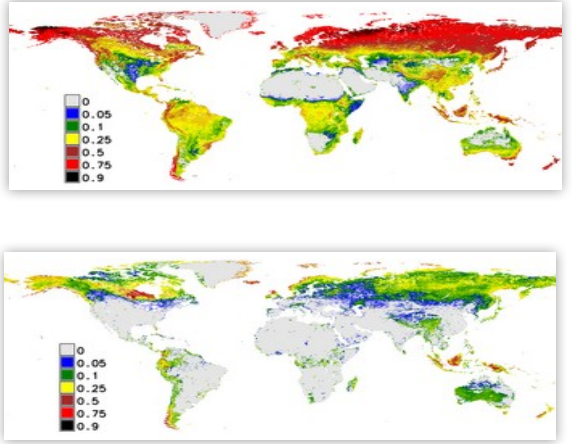
+

Soil bulk density (cg.cm^{-3})



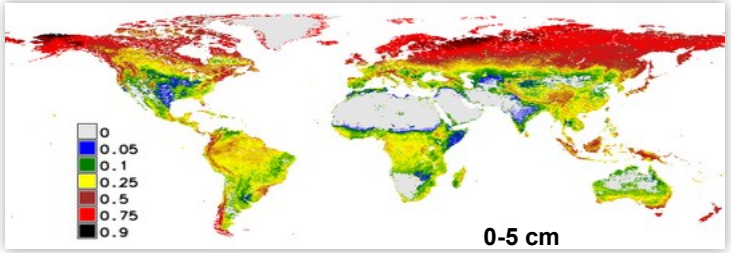
=

Organic matter volumetric fraction (m^3.m^{-3})



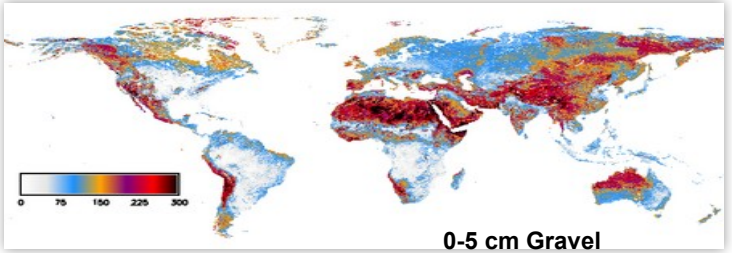
Ongoing work: revision of soil physical properties for ISBA

Organic matter volumetric fraction ($m^3.m^{-3}$)



+

Coarse fragments volumetric fraction ($m^3.m^{-3}$)



→ **New possibility for soil physics**

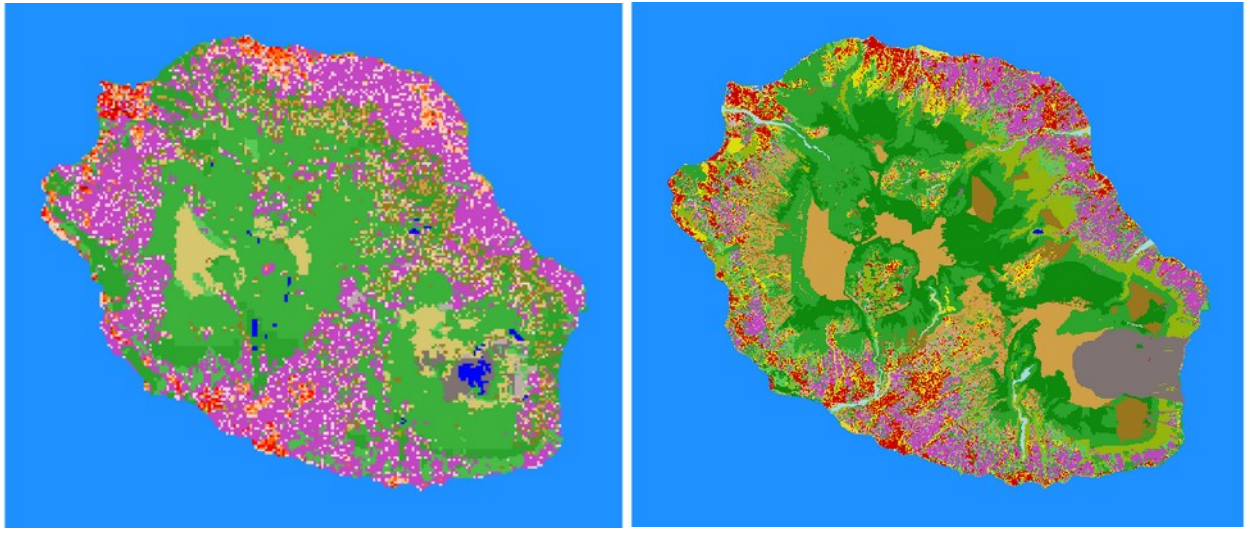
- ✓ Improvement of the Decharme et al. (2016) parameterization of the impacts of soil organic matter on soil hydrological and thermal properties
- ✓ Account for the presence of soil gravel (Mixed soil theory)

Ongoing work: ML/AI-based algorithm to retrieve HR land cover in the overseas

Value	Label		Color
1	Sea and oceans		
2	Lakes		
3	Rivers		
4	Bare soil	(NO)	
5	Bare rock	(ROCK)	
6	Permanent snow	(SNOW)	
7	Boreal broadleaf deciduous	(BOBD)	
8	Temperate broadleaf deciduous	(TEBD)	
9	Tropical broadleaf deciduous	(TRBD)	
10	Temperate broadleaf evergreen	(TEBE)	
11	Tropical broadleaf evergreen	(TRBE)	
12	Boreal needleleaf evergreen	(BONE)	
13	Temperate needleleaf evergreen	(TENE)	
14	Boreal needleleaf deciduous	(BOND)	
15	Shrubs	(SHRB)	
16	Boreal grassland	(BOGR)	
17	Temperate grassland	(GRAS)	
18	Tropical grassland	(TROG)	
19	Winter C3 crops	(C3W)	
20	Summer C3 crops	(C3S)	
21	C4 crops	(C4)	
22	Tree cover, flooded	(FLTR)	
23	Shrub or herbaceous cover, flooded	(FLGR)	
24	Urban LCZ1: compact high-rise	(CPHR)	
25	Urban LCZ2: compact midrise	(CPMR)	
26	Urban LCZ3: compact low-rise	(CPLR)	
27	Urban LCZ4: open high-rise	(OPHR)	
28	Urban LCZ5: open midrise	(OPMR)	
29	Urban LCZ6: open low-rise	(OPLR)	
30	Urban LCZ7: lightweight low-rise	(LWLR)	
31	Urban LCZ8: large low-rise	(LALR)	
32	Urban LCZ9: sparsely built	(SPAR)	
33	Urban LCZ10: heavy industry	(INDU)	

ECO-SG 2019 300m

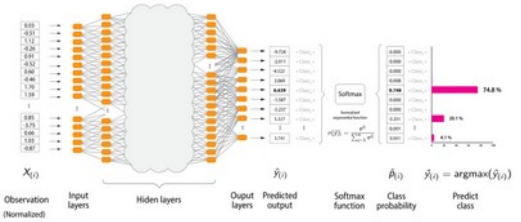
Ground Truth (Theia, Farce)



1) Sea and oceans
2) Lakes
3) Rivers
4) Bare soil
5) Bare rock
8) Temperate broadleaf deciduous
11) Tropical broadleaf evergreen
13) Temperate needleleaf evergreen
15) Shrubs
18) Tropical grassland
20) Summer C3 crops
21) C4 crops
22) Tree cover, flooded
23) Shrub or herbaceous cover, flooded
25) Urban LCZ2: compact midrise
26) Urban LCZ3: compact low-rise
27) Urban LCZ4: open high-rise
28) Urban LCZ5: open midrise
29) Urban LCZ6: open low-rise
30) Urban LCZ7: lightweight low-rise
31) Urban LCZ8: large low-rise
32) Urban LCZ9: sparsely built

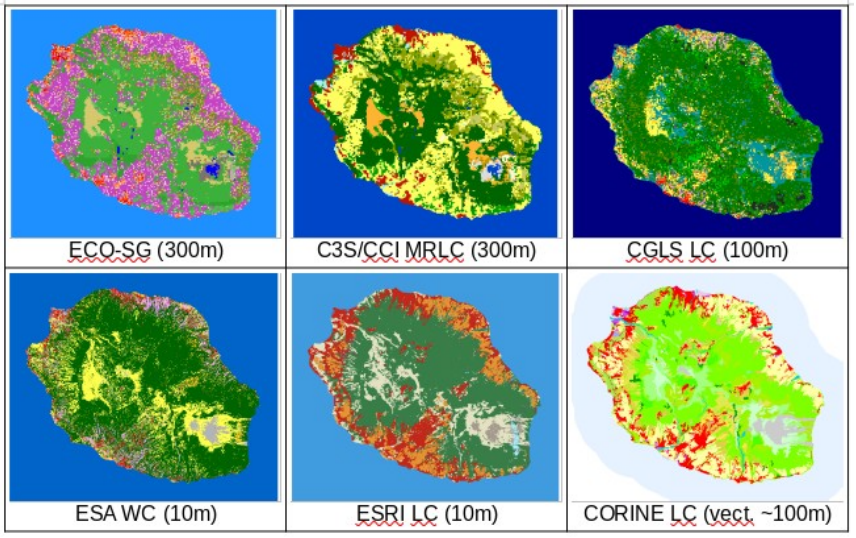
Reunion Island

Dense Neural Network

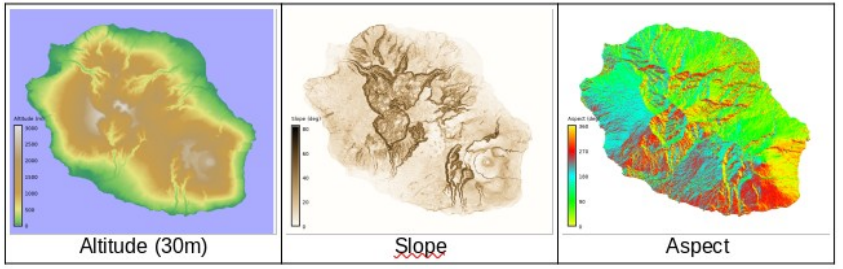


24 Predictors to train the DNN

✔ Land cover maps



✔ SRTM, 30m



✔ SOILGRIDS sand/clay 250m

✔ Global Forest Canopy Height, Maryland, 300m

✔ ECO-SG LAI, 300m

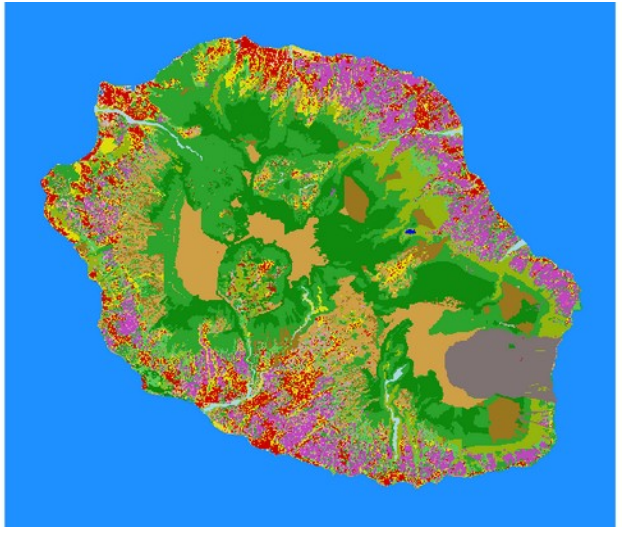
➔ Best results

- Corine LC water mask
- GeoClimate/OSM urban mask
- A specific DNN for the 34 nature classes
- Use of 24 predictors

Dense Neural Network

DNN map, 30m

Ground Truth



Summary

- **SURFEX** LSM for all MF's models
- Large variety of ancillary data used
- ECOCLIMAP-SG LC a specific algorithm for hectometric scale

Work in progress:

- To better represent urban parameters (GeoClimate/OSM)
- To improve soil hydrodynamic properties (SOM, gravels)
- To derive HR maps from ML-AI methods