Configuration e923 with PGD (version 2015)

Martina Tudor using many informations from Maria Derkova, Francoise Taillefer, Neva Pristov

October 2015

Abstract

This document describes the way the clim files, used by ISBA surface scheme, that contain "climatological" background fields can be created on beaufix in Meteo France. The 12 monthly files per domain can be produced in Meteo France by running the so called configuration e923 since the procedure requires the input databases available there. There are some features in the fields of climate files that are unrealistic and rather different from nature. See Abstract B for a brief overview of what are the fields in the clim file and what are possible problems with them. A number of figures included in this document illustrate what these features are. The unrealistic features in height of topography and land sea mask were cleared by the usage of the new database. But only these two fields are used from the new database, while the other fields are from the old database, so most of the features described in the Appendix B remained.

This report also describes one option how these features can be corrected in the climate files used for operational forecast using the physiographic input data from another database. This database is used for the SURFEX fields. First the SURFEX files are created and then these fields are used to compute the fields usually produced by e923. For most of the variables, the correction consists of computing the field using input from the SURFEX file. The computations are done a posteriori after the output file of a e923 Step is finished using tools that enable manipulation of the grid-point fields in the file.

This document also describes which scripts can be used to execute particular steps and the location of the script on ***beaufix*** computer in Meteo France. However, some of the computation executed on the fields to correct them were performed on the computer in Coratian Meteorological Service. The tool ***facalc*** and the *namelist*s can be provided/ported elsewhere. The clim files were created in different resolutions, varying from 15.4 km to 250 m using the tools described here.

1. Introduction

The clim files are used when running interpolation from one grid to another. The fields are interpolated as differences from “climate”. This is supposed to minimize the interpolation error. Other fields from the clim file are used to define surface properties in the model forecast.

The configuration e923 creates the climatological files, containing the following 2D grid point fields:
- constant fields describing orography,
- constant fields describing surface, soil, vegetation,
- monthly climatological values for soil or surface variables and
- monthly climatological fields for aerosols and ozone profiles.

Configuration e923 is used to create fields that contain so called climatological or constant fields. Some of these fields vary during the year and other are constant. The result are 12 monthly climatological files. The configuration e923 is executed in 8 steps, numbered 1 to 9, while the step number 7 is omitted in the standard procedure (it is related to the aqua planet simulations). A climatological file is created in the following 8 steps (the skipped step is also listed):

step 1 - definition of numerous fields describing orography and land-sea mask, depending on the *namelist*, few fields that describe topography can be created from a PGD file (using one database – the same as for SURFEX) while other fields are computed from the usual database (another one), creates one output file
step 2 - definition of surface, soil and vegetation characteristics without annual cycle, creates one output file, uses fields computed in the previous step and input,
step 3 - definition of monthly climatological values of soil temperature and moisture, modification of albedo and emissivity according to the climatology sea-ice limit, creates 12 output files, uses fields computed in the first two steps and input,
step 4 - definition of the vegetation characteristics and modification of several surface characteristics, uses fields computed in the first three steps and input,
step 5 - modification of fields created by step 2 or 4 over land from high resolution datasets for each month, uses fields computed in steps 1, 2 and 4 and input,
step 6 - modification of climatological values for soil temperature and moisture using “new” climatology, uses fields computed in steps 1 and 3 and input,
step 7 – modification of fields over the water surfaces (not run for usual e923)
step 8 – computes monthly fields of A, B and C coefficients for profiles of ozone
step 9 – computes monthly fields of aerosols (sea, land, soot, desert).

Steps 4 to 9 have to be run separately for each monthly file. Each step uses some of the fields created in previous steps as input. Additionally, there is Step 10, that is also run for the aqua planet and not used for common clim files.

2 Input data

Each step of the configuration e923 uses a number of input files that contain different fields in different resolutions, some of the input files cover the whole globe, while other cover a smaller area and most of them use the output fields from the previous step. The default values for the covered area and resolution of the input files are set in incli0.F90, these default values are listed in Table 1, but can be changed in the *namelist* NAMCLI (see Table 2). The names of the required input files for each step are defined in the model code and these will be described in more details further in the text (see Section 3).

Table 1. Parameters change according to N923 value (step of 923). Finally, the values for NGLOBX/Y are set equal to NDATX/Y, and the resolution of input data EDLON/LAT is computed from NDATX/Y for datasets 1-4, 6, 8 and 9. The values for step 6 are modified in the *namelist*. The rest are using these default values.

N923	1	2	3	4	5	6	8	9
LIEEE	T	T	T	T	T	T	F	F
NDATX	8640	360	432	360	860	240	144	72
NDATY	4320	180	216	180	420	120	73	45
coverage	global	global	global	global	SW(-25,30)	global	global	global
Resolution (degrees)	0.042	1	0.83	1	0.1	1.5	2.5	5

The clim files are created by executing computations related to the 8 steps. Due to limitations in the spectral transform of topography, it is possible to run configuration e923 using one processor only. Each step requires input files named in a prescribed way (the names are hardcoded in the subroutines) so the input files are copied from the source directories to the working directory with the required name before executing computations.

Some steps would work (well, execute something) even without the input files (using only output from the previous step) and write the output file. Such steps simply check for the existence of the input files and execute particular computations only if the input file exists. The only indication of the missing input files would be in the NODE file. So some variables change and others do not, depending on the existence of the input files. Currently there are several high resolution files missing/not used in different steps. See subsections of Section 3 for more details.

The computations related to all steps can be done using a single script. The only thing that changes is the *namelist* and mostly in a minimum way since each step requires different N923 in the *namelist*. Most of the steps use the default definitions of the input files (their geometry and coverage is as listed in Table 1). The only exception is Step 6, where an alternative data set is used for input and the resolution of the input increased from 1.5 to 1 degree (longitude and latitude).

Table 2. Parameters in *namelist* NAMCLI (YOMCLI) that define the input file: format, coverage and resolution as well as several parameters used in various steps of e923.

Parameter	Default	Meaning
LIEEE	T	Format of input file, T if ieee format is used, ASCII otherwise
LGLOBE	T	Coverage of input data set, T if global dataset
NPINT	1	size of the interpolation box ; >= minimum computed by (E)INTER0
NDATX	1	x-size of the dataset (longitude)
NDATY	1	y-size of the dataset (latitude)
ELONSW	0	Longitude and Latitude of the southwest and Northeast corners of local dataset used in eincli5
ELATSW	-90	
ELONNE	360	
ELATNE	90	
EDLON	360	resolution of the dataset (in degrees)
EDLAT	180	
NGLOBX	1	corresponding numbers of points on the globe
NGLOBY	1	
SVEG	0.02	threshold for significant vegetation cover, it is set to zero if not LSOLV
LZ0THER	T	.FALSE. if no orographic part in the thermal roughness length, should be combined with LZ0HSREL in the subsequent forecast run!!!
SFCZ0	1	scaling factor for the secondary part of z0 (urban., veget.)
RSTR	5	threshold for temperature-RTT in presence of snow (eincli6)
RSWR	0.02	threshold for mean moisture in presence of snow eincli6
NSLICE	1	number of packet used to slice the domain (LAM only)

The defaults in Table 2 change depending on the step in the configuration e923 (the parameter N923 in the *namelist* NAMMCC), see Table 1 for details. Particular attention should be put to the LZ0THER parameter. The old clim files were computed using **LZ0THER=T** (Table 2) and FACZ0=0.53 and NLISSZ=3 (Table 6). The old scheme computes surface turbulent fluxes of heat and moisture using the surface roughness length for heat that is computed with the contribution of sub-grid orography so e001 is run with **LZ0HSREL=F**. But the new files are computed using **LZ0THER=F**, FACZ0=1 and NLISSZ=1. Therefore, the new formulation means longer roughness length for momentum but much smaller roughness length for heat and moisture. So with the new field, the surface turbulent fluxes for heat and moisture are computed without the contribution of the sub-grid scale orography and the forecast run using new clim file should be run with **LZ0HSREL=T** in the *namelist*.

2.1 Input topography from a PGD file

The novelty in e923 is the usage of PGD. This allows using new topography from the new database. A PGD file is created using a separate script that has to be run before the step 1. Using the topography and land sea mask from the new file is invoked by setting

LNORO=T in the *namelist* NAMCLA. If one wants to use new (alternative) land sea mask (LNLSM=T), one also needs to use the new topography file. Actually, in the code, the alternative file containing the land sea mask and topography is read if LNORO. If PGD topography is imported (LIPGD=T) in the Step 1 of e923, envelope is added to the topography using multiplication factor FENVN and the topography variance from the old database (different than the database used for the PGD file).

The clim files are created as a combination from several databases. As a consequence land-sea mask and height of topography are taken from one database (used for SURFEX), but proportion of land, standard deviation of orography and other parameters describing topography are taken from another database.

A script is run that creates a PGD file containing surface parameters in a format suitable for subsequent use in the configuration 923. The topography and land sea mask from this file will be read in Step 1 of the configuration 923 if LNORO (and LNLSM). An example of this script is ***make_pgd_923_model_hr88*** in ***/home/gmap/mrpm/tudorm/scr*** on ***beaufix.*** In the beginning one defines the input data used to create the PGD file. These data files are in the directory:
DIR_DATA=/home/gmap/mrpa/taillef erf/SFX_databases
Each data file also has a header file, so both have to be linked and the names of the links are defined in the *namelist*, these are for the files containing data on clay and sand:
ln -s ${DIR_DATA}/CLAY_HWSD_MOY.dir clay.dir
ln -s ${DIR_DATA}/CLAY_HWSD_MOY.hdr clay.hdr
ln -s ${DIR_DATA}/SAND_HWSD_MOY.dir sand.dir
ln -s ${DIR_DATA}/SAND_HWSD_MOY.hdr sand.hdr

Here a link to the ecoclimap database is made
ln -s ${DIR_DATA}/ECOCLIMAP_I_GLOBAL_V1.6.dir ecoclimap.dir
ln -s ${DIR_DATA}/ECOCLIMAP_I_GLOBAL.hdr ecoclimap.hdr

And here a link to the file containing topography (and land sea mask) in 30'' resolution:
ln -s ${DIR_DATA}/GMTED2010_30.EHdr.dir orog.dir
ln -s ${DIR_DATA}/GMTED2010_30.EHdr.hdr orog.hdr

however, there is even a higher resolution alternative for the topography (so yes, the new high resolution topography and land sea mask already have an alternative in even higher resolution):
ln -s ${DIR_DATA}/GMTED2010_075.EHdr.dir orog.dir
ln -s ${DIR_DATA}/GMTED2010_075.EHdr.hdr orog.hdr

that contains data in 7.5'' resolution. The clim files in 2 km resolution (and higher) were created using this high resolution input while clim files in lower resolution (4 – 15.4 km) used input from the lower resolution input topography.

#These files are necessary files to run surfex-pgd
ln -s ${DIR_DATA}/ecoclimapI_covers_param.bin ecoclimapI_covers_param.bin
ln -s ${DIR_DATA}/ecoclimapII_eu_covers_param.bin ecoclimapII_eu_covers_param.bin

The example script listed above also contains the *namelist*, the explanation of several parameters is given in Table 3.

Table 3. *Namelist* parameters needed to create the PGD file (one should modify only the

parameters in bold font that define the domain grid).

Namelist	Parameter	Value/meaning
NAM_IO_OFFLINE	CSURF_FILETYPE	'FA '
	CPGDFILE	PGD_ALD_923 name of output file
NAM_PGD_GRID	CGRID	CONF PROJ type of projection
NAM_CONF_PROJ	**XLAT0**	**Reference latitude in degrees**
	XLON0	**Reference latitude in degrees**
	XRPK	**sin(XLAT0)**
	XBETA	0
NAM_CONF_PROJ_GRID	**XLONCEN**	**Longitude of centre (equal to XLON0)**
	XLATCEN	**Latitude of centre (equal to XLAT0)**
	NIMAX	**Number of grid points in the X direction of the C+I zone (without the extension zone)**
	NJMAX	**Number of grid points in the Y direction of the C+I zone (without the extension zone)**
	XDX	**Grid step in x direction in meters**
	XDY	**Grid step in y direction in meters**
NAM_FRAC	LECOCLIMAP	T use ecoclimap database
NAM_COVER	YCOVER	'ecoclimap' name of file containing data
	YFILETYPE	DIRECT type of access
NAM_ISBA	YCLAY	'clay' name of file containing data for clay
	YSAND	'sand' name of file containing data for sand
	YCLAYFILETYPE	DIRECT type of access
	YSANDFILETYPE	DIRECT type of access
NAM_ZS	YZS	'orog' name of file containing data for topography
	YFILETYPE	DIRECT type of access
NAM_ZS_FILTER	NZSFILTER	0, number of times a filter is applied to the topography, zero for no filtering
NAM_PGD_ARRANGE_COVER	LTOWN_TO_ROCK	T if one converts towns to rocks
NAM_PGD_SCHEMES	CTOWN	'NONE' type of town scheme

Table 4. Names of fields in the FA 923 PGD file, the range of values these fields assume on a LACE coupling domain (the range for topography fields depends on the domain and resolution) and what they represent.

Field	Range	meaning
S1D_FRAC_SEA	0-1	Fraction of the grid cell occupied by sea
S1D_FRAC_NATURE	0-1	Fraction of the grid cell occupied by natural areas
S1D_FRAC_WATER	0-1	Fraction of the grid cell occupied by water surface (but not sea surface)

S1D_FRAC_TOWN	0-1	Fraction of the grid cell occupied by urban areas (equal to 0 in the output file)
S1D_COVER001	0-1	Land sea mask
S1D_COVER002	0-1	lakes
S1D_COVER003	0-1	rivers
S1D_COVER004	0-1	Bare land
S1D_COVER005	0-1	Rocks
S1D_COVER006	0-1	Permanent snow (glaciers?)
S1D_COVER007	0-1	Urban areas
S1D_ZS	-4 - 3230	Height of topography in meters (range depends on the domain and resolution)
S1D_AVG_ZS	-4.3-3391	Average height of topography
S1D_SIL_ZS		Silhouette topography
S1D_SSO_STDEV	0-943.9	Standard deviation of topography (10^{20} above sea)
S1D_MIN_ZS		Minimum sub-grid topography
S1D_MAX_ZS		Maximum sub-grid topography
S1D_SSO_ANIS	0-1	Anisotropy of topography (10^{20} above sea)
S1D_SSO_DIR	-1.57-1.57	Direction (10^{20} above sea)
S1D_SSO_SLOPE	0-0.529	Slope of topography (10^{20} above sea)
S1D_CLAY	0-1	Fraction of the grid cell occupied by clay soil type (10^{20} above sea)
S1D_SAND	0-1	Fraction of the grid cell occupied by sand soil type (10^{20} above sea)

The script creates the output file named PGD_ALD_923.fa (the name is defined in the *namelist,* and the extension .fa is added in the end). The output file has to be stored somewhere, for example in the $WORKDIR. This output file is in the FA format, it can be examined using popular tools such as *frodo* and *edf.* However, the contents are rather different from the usual substance of the FA files that aladiners (using ISBA) are used to. First of all, the procedure that creates this PGD file assigns the truncation values for NSMAX and NMSMAX as if the data are in linear grid (this can be a problem when one wants to have a clim file for quadratic grid). The file contains a number of fields and many parameters that are actually not fields that cover the whole domain. When the file is read by frodo, these parameters are listed in the section with other fields. Some fields interesting for further use are listed in Table 4.

This file, PGD_ALD_923.fa, is later used in the Step 1 of the configuration e923 as a source of alternative topography and land sea mask data. Unfortunately, as will be explained later, only these two fields are taken from the new, high resolution database in 30" resolution (or even higher resolution database in 7.5" resolution), and the other fields are taken from the old database. This makes the final surface fields computed by the configuration 923 inconsistent and is expected to have an impact on the model forecast (especially if one looks at some long term budgets). However, the model forecast will probably run with almost anything put into the clim file (with the exception of LZ0THER that did cause and abortion in the model forecast when combined with wrong LZ0HSREL). Once this PGD file is created, one can proceed to the Step 1 of e923. However, one can create yet another PGD file that would be a SURFEX type file, in LFI format, with the same content, that will be used subsequently to correct the output of e923 using more fields from the new database.

2.2 Make another PGD file, "SURFEX type" file

Using another, but very similar script (an example *make_pgd_surfex_hr88* can be found in the same directory as the previous script), and *namelist* contained in it, one can create a SURFEX like lfi file containing the variables that describe the surface climatology. The procedure is very similar to the previous one, so only differences are listed here:
- the output fields have a prefix "SFX." instead of "S1D_" but the same content as corresponding field in the previously created FA file.
- the input files should be the same as for the previous step!
- CSURF_FILETYPE='LFI   ' in NAM_IO_OFFLINE defines that the output file is in the LFI format.

The contents of this file are subsequently (at the end of Step 1) written into an FA file with correct grid size (that includes an extension zone) and truncation (as wished) assigned by the another input file in FA format. One should take care that the geometry of the lfi file and the FA file fit so that the contents of the lfi file fill the C+I zone of the fa file. When the first FA file (created as described in the subsection 2.1) is used as input, there was a problem with the truncation for the clim file with quadratic truncation.

3. Steps of e923

3.1 Step 1

Step 1 of the configuration e923 computes 9 fixed fields that describe the orography. The input datasets were prepared from GLOBE25 or from GTOPT030 data in resolution 2'30", the files contained data describing the global orography in 2'30" (0.042 degrees) resolution (4.67 km at the equator). These fields are computed from US NAVY 10' and NOAA 30" datasets. The names of the input files are hard-coded in the subroutine eincli1.F90. Input files listed in Table 5 are taken from the directory:
/home/gmap/mrpa/tailleferf/923/RELIEF_G/GTOPT030/
and then also take the new topography from the PGD file
cp $WORKDIR/PGD_ALD_923_HR88.fa   Neworog

Table 5. Files used as input for the Step 1. The aerial coverage and resolution of files used in each step of configuration e923 is listed in Table 2. These files are referred to as the old database in the text.

Name of file	contents
Water_Percentage	Percentage of water surfaces, 0 for land, 100 for water
Oro_Mean	Mean height of orography in meters
Sigma	Subgrid standard deviation of the mean height in meters
Nb_Peaks	Number of subgrid peaks
Urbanisation	Fraction of grid occupied by urban areas
Dh_over_Dx_Dh_over_Dy	Component of orography variance tensor
Dh_over_Dx_square	Component of orography variance tensor
Dh_over_Dy_square	Component of orography variance tensor
Hmax-HxH-Hmin_ov4	$(H_{max}-H_{mean})*(H_{mean}-H_{min})/4$
Oro_Max	Maximum height of topography (not used in Step 1)
Oro_Min	Minimum height of topography (not used in Step 1)

The fields computed by the Step 1 are:

1) SURFIND.TERREMER – land sea mask, can have only 2 values, values over land are fixed to 1, and 0 for water surfaces if the percentage of water surface over the area of the grid cell is greater than the SMASK parameter that can be controlled via *namelist* NAMCLI, and was set to 0.5. The “old” land sea mask is applied to correct the roughness length and compute the fraction of land, and the new percentage of water is read from the new file afterwards and used to compute the new land sea mask (see Figure 1 for example). But the contents of this new file is not used for the field describing the percentage of water (or land) in the grid cell nor to correct the fields that should have certain default values above the sea surface.

Table 6. Parameters set in NAMCLA (module YOMCLA) used to tune the representation of orography, their default values and meaning. One should keep in mind that there are other important parameters defined in NAMCLI (Table 3) and other *namelist*s (Appendix A) and the final result depends on the combination of factors.

Parameter	Default	Meaning
FENVN	0	Factor for the envelope orography on pole of interest.
FENVS	0	Factor for the envelope orography on the antipode of the pole of interest
LNORO	F	Key for reading a new orography on a separate file
LNLSM	F	Key for reading a new land-sea mask on a separate file
LIPGD	F	Key to take a new GP orography from a FA PGD file
LKEYF	T	Key for spectral fit on the orography, used only for LAM (this part run only if not LNORO)
NLISSR	0	Number of calls to ELISLAP to smooth the orography (LAM only)
LNEWORO	F	Switch to the Bouteloup cost function
LNEWORO2	F	Switch to the Jerczynski cost function
LSPSMORO	F	Switch for spectral smoothing of orography by a diffusion operator (linear grid)
QMAX	4	Maximum value of weight
QMIN	2	Minimum value of weight
HMIN	150	Reference height in the formulation of weight
HDIM	1	Scaling factor for the orography deviation
XINCOC	0	Part of the fraction of sea in the weight
SCEXT	0	Scaling factor for the weight of the extension zone
QPOWER	3.5	Jerczynski cost function exponent in the additional term
QCONST	0.4	Jerczynski cost function scaling factor in the additional term
FACE	1	Jerczynski cost function 1. for initial formulation , 0. for polynomial one
NLISSP	0	type of the spectral smoothing of orography; 0 : no, 1 : importing a lower resolution one , 2 : in optimization
FLISB	0	tuning parameter if NLISSP=2
FLISA	0	threshold in spectral space if NLISSP=2
NLISSZ	3	Number of calls to LISLAP to smooth the roughness length
FACZ0	1	Scaling factor for the orographic part of Z0

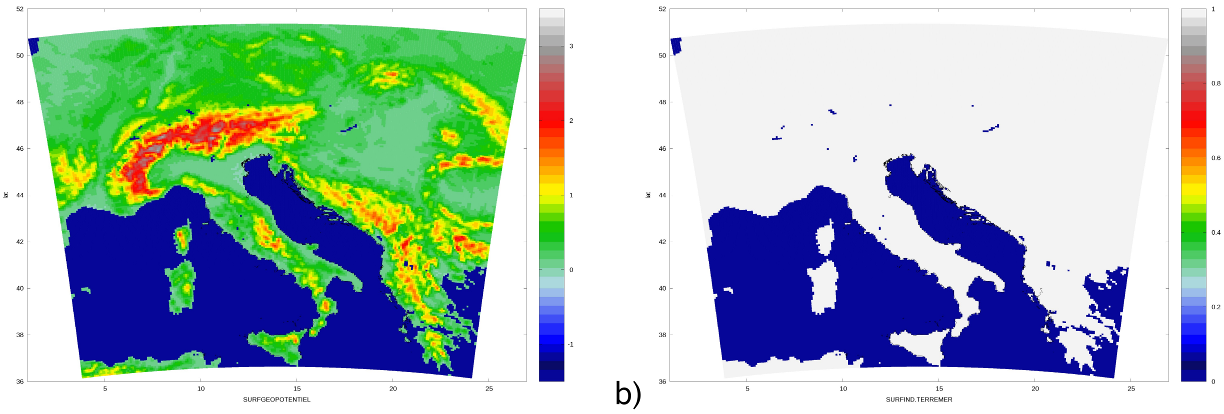
2) SURFGEOPOTENTIEL – terrain height as geopotential (height in meters multiplied by 9.81 m/s2) as a gridpoint field and SPECSURFGEOPOTEN as its spectral counterpart. This field can include the effect of envelope, controlled by the parameter FENVN in *namelist* NAMCLA, and the height of surface is computed according to:

$$Hs=H_{mean}+FENVN*Hstdev*(1-P_{water})$$

where P_{water} is the percentage of water in the grid cell (the old one!) and Hstdev is computed as described below.

However, the envelope is also added if LIPGD (if we use the new height of topography from the new database for input) using the same formula, therefore if we import the new topography from a PGD file, but the standard deviation of topography (Hstdev) used in the above formula is from the old database. **The results shown here were computed with FENVN=0,** therefore there was no envelope added to the mean orography.

These first two fields (land sea mask and height of topography) are subsequently read again from a file that contains the new (alternative) topography named *Neworog* created as a PGD file in the FA format. The topography is subjected to spectral fit (if LKEYF) and both grid-point and spectral fields are multiplied by g=9.81 before being written to the output file (see Figure 1 for an example of topography and land sea mask). Usage of new topography finally cleared some issues with false peninsulas along the Italian coastline.



a) b)

Figure 1. Surface geopotentiel (a) (divided by 9.81 and 1000 to represent terrain height in km and forced to -1 above the sea) and land sea mask (b) with LNORO, and LNLSM true for the domain in 8 km resolution.

3) SURFET.GEOPOTENT is the standard deviation of sub-grid topography multiplied by g=9.81m/s2. This field is taken from the old database by the standard Step 1 of the configuration e923. It is computed as:

$$Hstdev=sqrt(\sigma_H^2+\sigma_h^2+(H_{max}-H)(H-H_{min})/4)$$

where σ_H = avgint(H^2)-H^2 is the terrain variance resolved in the input file but not in the output topography (rather questionable what this is once the output resolution approaches the resolution of the input file), σ_h is the unresolved variance read from the *Sigma* input file, and $(H_{max}-H)(H-H_{min})/4$ is simply the contents of file *Hmax-HxH-Hmin_ov4*. The function avgint is actually computed by einter1 subroutine where the data are interpolated to the closes point and averaged using a neighbourhood of points defined by NPINT (NPINT=1 was used here).

4) SURFVAR.GEOP.ANI is the anisotropy coefficient, it is equal to 1 for isotropic surfaces (sea) and varies from 0 to 1, where lower values mean that the terrain is particularly varying in one direction, but not in another. This field is computed using the fields taken from the old database by the standard Step 1 of the configuration e923. The output of

eganiso is the square of the anisotropy coefficient. But in acdrag it is used as anisotropy (not the square value). The anisotropy gamma and direction theta are computed as

$$\gamma^2 = \frac{P_1 + P_2 - \sqrt{(P_1 - P_2)^2 + 4P_3^2}}{P_1 + P_2 + \sqrt{(P_1 - P_2)^2 + 4P_3^2}}$$

$$\theta = atan\left(\frac{-(P_1 - P_2) + \sqrt{(P_1 - P_2)^2 + 4P_3^2}}{-2P_3}\right) - atan(\frac{gnordl}{gnordm})$$

$$P_1 = \frac{\left(\frac{\partial H}{\partial x}\right)^2 \sigma_H^2}{\left(\frac{\partial H}{\partial x}\right)^2 + \left(\frac{\partial H}{\partial y}\right)^2} + \frac{\left(\frac{\partial h}{\partial x}\right)^2 \sigma_h^2}{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}$$

$$P_2 = \frac{\left(\frac{\partial H}{\partial y}\right)^2 \sigma_H^2}{\left(\frac{\partial H}{\partial x}\right)^2 + \left(\frac{\partial H}{\partial y}\right)^2} + \frac{\left(\frac{\partial h}{\partial y}\right)^2 \sigma_h^2}{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}$$

$$P_3 = \frac{\frac{\partial H}{\partial x}\frac{\partial H}{\partial y}\sigma_H^2}{\left(\frac{\partial H}{\partial x}\right)^2 + \left(\frac{\partial H}{\partial y}\right)^2} + \frac{\frac{\partial h}{\partial x}\frac{\partial h}{\partial y}\sigma_h^2}{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}$$

5) SURFVAR.GEOP.DIR is the direction of the principal axis of topography (in radian). This field is computed using the fields taken from the old database by the standard Step 1 of the configuration e923. The angle of topography is computed from the components of the tensor and then the angle of the grid orientation is added so the final angle theta is not restricted to the range from -pi/2 to pi/2.

6) SURFZ0REL.FOIS.G is the roughness length of the bare surface multiplied by g=9.81m/s2, it can be scaled using an arbitrary parameter FACZ0 in the *namelist* NAMCLA, the experiments shown here used FACZ0=1. This field is computed using the fields taken from the old database by the standard Step 1 of the configuration e923. In e923, the surface roughness is computed as a combination of standard deviation and square root of the density of the number of peaks from both resolved and unresolved topography:

$Z_{0ter}=\sigma_H sqrt(N/S)+\sigma_h sqrt(n/S)$

where N is the number of peaks and σ_H = avgint(H^2)-H^2 is the standard deviation resolved in the input file but not in the output topography, n is the number of peaks in the NB_peaks input file, σ_h is the unresolved standard deviation read from the *Sigma* input file, S is the surface of the grid cell.

7) At the end of the Step 1, the momentum roughness length SURFZ0.FOIS.G is equal to the z0 for bare land. It is subsequently corrected using the vegetation roughness, but the final field is dominated by the contribution of the bare land and contains the same unnatural features described in Section 3.2.1.

8) SURFPROP.TERRE is the field that describes the fraction of land in the gridbox. It is equal to 1 for purely land points, 0 for the sea points, and varies from 0 to 1 over the coastlines, land with small lakes and rivers. It should match the land sea mask field, but in the Step 1 it is taken from the old database, even if the land sea mask is taken from the new one. This field can have substantially different values. The proportion of land is used in the subsequent steps of the configuration e923. The values should be computed from

the new database and put into the Const.clim file before the subsequent steps of the configuration e923.

Step 1 is actually performed in two steps. First Step 1a with quadratic grid truncation and then Step1b with linear grid truncation. The results presented here are for quadratic grid truncation. This affects only the representation of surface geopotential in the spectral space and the same features appear in the grid point fields.

3.2.1 Inserting fields from the new database to the clim file (after Step 1 is finished)

As mentioned before, only two of the fields describing topography are taken from the new database, and other fields are taken from the old one. Since the proportion of land had several strange features over the Adriatic after the Step 1 and not corrected in the subsequent steps so these features remained in the final clim files (Figure 2a, not-existent peninsulas along Italian coastline, rather messy Croatian coastline).

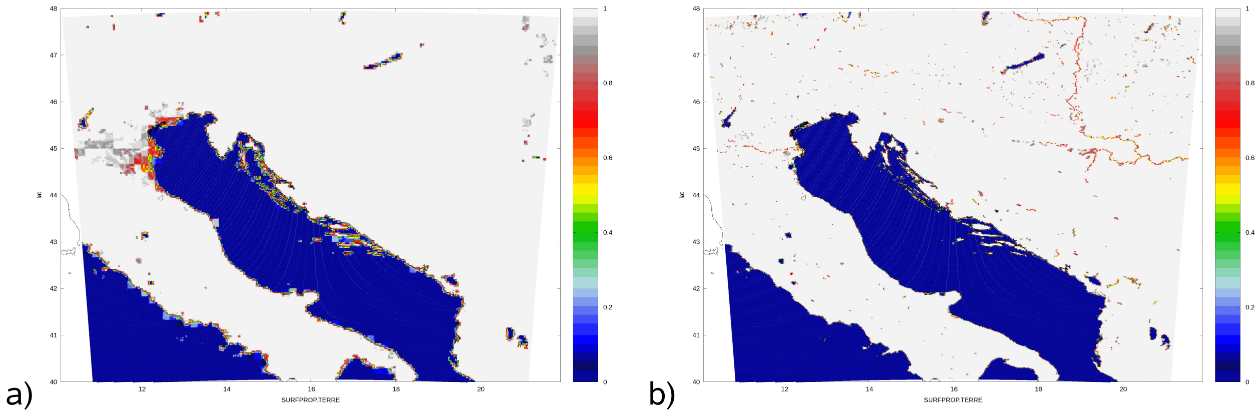


Figure 2. Proportion of land after the Step1 (a) and computed from the SURFEX PGD file (b). These figures show data for 2 km resolution domain.

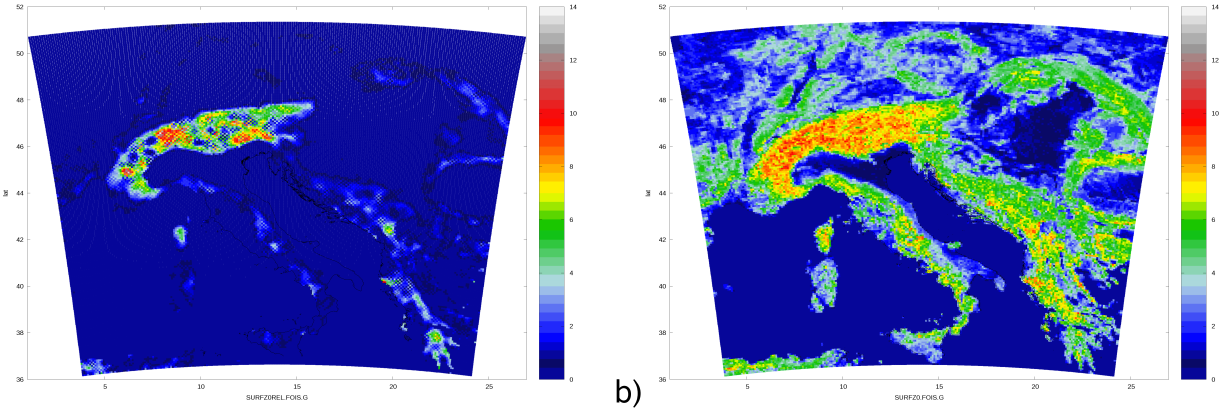
Therefore an FA file was created from the SURFEX PGD file that had the same geometry as the Const.clim file and then the percentage of land was computed as:

$$P_{land}=1-P_{sea}-P_{water}$$

where P_{sea} and P_{water} are SFX.FRAC_SEA and SFX.FRAC_WATER. The resulting field for the percentage of land is shown in Figure 2b. The resulting field looks closer to what we can find in nature (or GoogleEarth). This field was inserted into the Const.clim file overwriting the old SURFPROP.TERRE field.

Furthermore, surface roughness for the bare land exhibited a chessboard pattern over the Alps and other mountains (Figure 3a, this feature became apparent only when the value of the roughness length was plotted as a small square on the position of the grid point), this was considered unnatural and could affect the forecast quality there.

As a first guess, this parameter was taken from the new database as the square root of the standard deviation of topography times g and the resulting field is shown in Figure 3b. One can immediately see much higher values above the mountains such as the Alps, but even more so above the mountains along the Croatian coastline that were rather smooth in the old clim files. The z0 field can be tuned together with the turbulence scheme. The Figures 3 a and b show the roughness length in the clim file from the old database and the new field, the range of values in the new file is the same as the range in the old files.



a) b)

Figure 3. Surface roughness (a) and corrected field (b) in 8 km resolution (the values in the figures were divided by 9.81).

The standard deviation of topography in the SURFEX PGD file has values 10^{20} above the sea surface, this was first corrected to be the default value above the sea SZZ0M for purely sea points, since this is the value used in e923. If there is some sub-grid land, and the values were in reasonable range (therefore not 10^{20}), the value was kept since sub-grid islands do contribute to the roughness length. However, in e923, the values are set to SZZ0M if the proportion of land is less than SMASK. Since Croatia has several meteorological stations situated on islands that are sub-grid in rather high resolution, the effect of this can be tested. The same procedure was applied to the SURFZ0.FOIS.G and SURFZ0REL.FOIS.G fields since both have the same values at the end of the Step 1. The field describing standard deviation of topography in the clim file was also corrected.

Average height of topography and its standard deviation represents all the sub-grid topography, including islands where land occupies less than half of the grid cell, so the land sea mask is zero at these points. This means that we have an elevated water surface in the model in cases of isolated high islands, this is difficult to detect due to spectral representation that also produces waves in height of topography over the sea surface.

In order to harmonize the fields in the clim file so that all of them are taken from the same database, fields describing the anisotropy of topography and angle to the main axis were taken from the SURFEX PGD file. There, both of these fields had default values of 10^{20} above the sea, so this was first corrected to the default values used in the clim files, then the angle of topography was scaled since it varied from -90 to 90 (it is in degrees in the SURFEX PGD file). The anisotropy gamma and the angle theta are computed in SURFEX as:

$$K = \frac{1}{2}\left(\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2\right) \qquad L = \frac{1}{2}\left(\left(\frac{\partial h}{\partial x}\right)^2 - \left(\frac{\partial h}{\partial y}\right)^2\right) \qquad M = \frac{\partial h}{\partial x}\frac{\partial h}{\partial y}$$

$$\theta = \frac{1}{2} atan\left(\frac{M}{L}\right) \qquad \gamma = \sqrt{\frac{K - \sqrt{L^2 + M^2}}{K + \sqrt{L^2 + M^2}}}$$

One can immediately see that the output is gamma, not the square of gamma as in e923. The anisotropy from e923 is shown in Figure 4a and from SURFEX in Figure 4b. Larger values means that the unresolved topography is more isotropic. Low values indicate that terrain is changing dominantly in one direction and the values above the sea should be 1 (isotropic surface). Since gamma varies from 0 to 1, square of this field should yield lower values. This is why gamma has higher values in Figure 4b than in 4a.

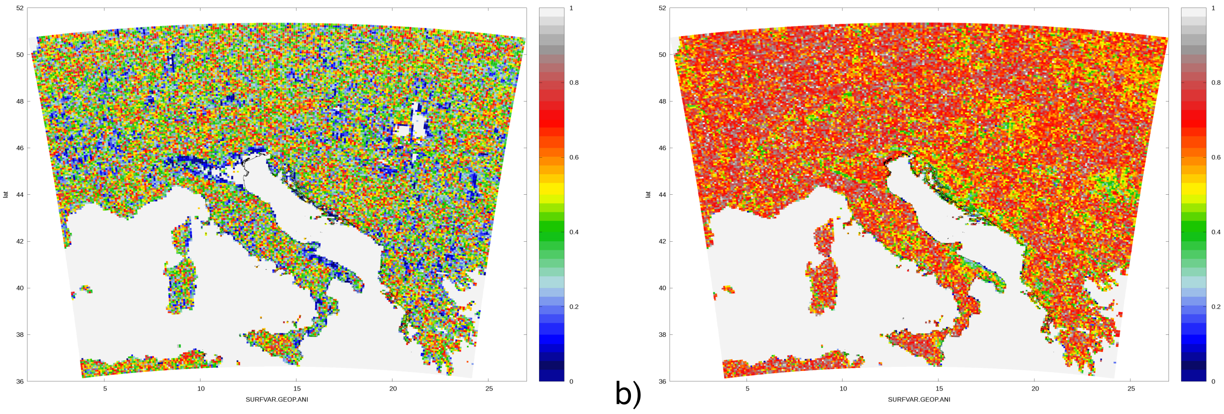


Figure 4. Anisotropy of unresolved topography in 8 km resolution, when computed by e923 from the old database (a) and taken from SURFEX (b).

In SURFEX theta is (according to the available documentation) defined as an angle to the x axis, positive northward and varies from -90 to 90. In code, theta is multiplied by 180/pi. Depending on the sign of L, the above theta is modified so that 90 degrees are added if L is negative and in the end, if the resulting angle is larger than 90 degrees, theta is reduced by 180 degrees. The link between the two thetas (the one from e923 and another from SURFEX) is less transparent. Let's assume that one can identify the following:

$$K = \frac{1}{2}(P_1 + P_2) \qquad L = \frac{1}{2}(P_1 - P_2) \qquad M = P_3$$

Then we start from the computations of theta in e923, how it is computed in eganiso, and insert the relations as defined above.

$$\theta = atan(\frac{-(P_1 - P_2) + \sqrt{(P_1 - P_2)^2 + 4P_3^2}}{-2P_3})$$

$$\theta = atan(\frac{-2L + \sqrt{4L^2 + 4M^2}}{-2M}) = atan(\frac{L - \sqrt{L^2 + M^2}}{M})$$

Since in SURFEX the angle theta is defined as one half of the arcus tangens function, one has to go to some basic trigonometic transformations. It is valid that:

$$atan(x) = \frac{1}{2}atan\frac{2x}{1 - x^2} \qquad for \qquad |x| < 1$$

$$atan(x) = \frac{\pi}{2} + \frac{1}{2}atan\frac{2x}{1 - x^2} \qquad for \qquad x > 1$$

$$atan(x) = -\frac{\pi}{2} + \frac{1}{2}atan\frac{2x}{1 - x^2} \qquad for \qquad x < -1$$

Inserting this property (well the first line) into the above formula for theta, one gets:

$$\theta = \frac{1}{2} atan \frac{2\frac{L-\sqrt{L^2+M^2}}{M}}{1-\left(\frac{L-\sqrt{L^2+M^2}}{M}\right)^2} = \frac{1}{2} atan \frac{2\frac{L-\sqrt{L^2+M^2}}{M}}{\frac{M^2-(L^2-2L\sqrt{L^2+M^2}+L^2+M^2)}{M^2}}$$

$$\theta = \frac{1}{2} atan \frac{2M(L-\sqrt{L^2+M^2})}{-2L^2+2L\sqrt{L^2+M^2}} = \frac{1}{2} atan \frac{2M(L-\sqrt{L^2+M^2})}{-2L(L-\sqrt{L^2+M^2})}$$

$$\theta = -\frac{1}{2} atan \frac{M}{L}$$

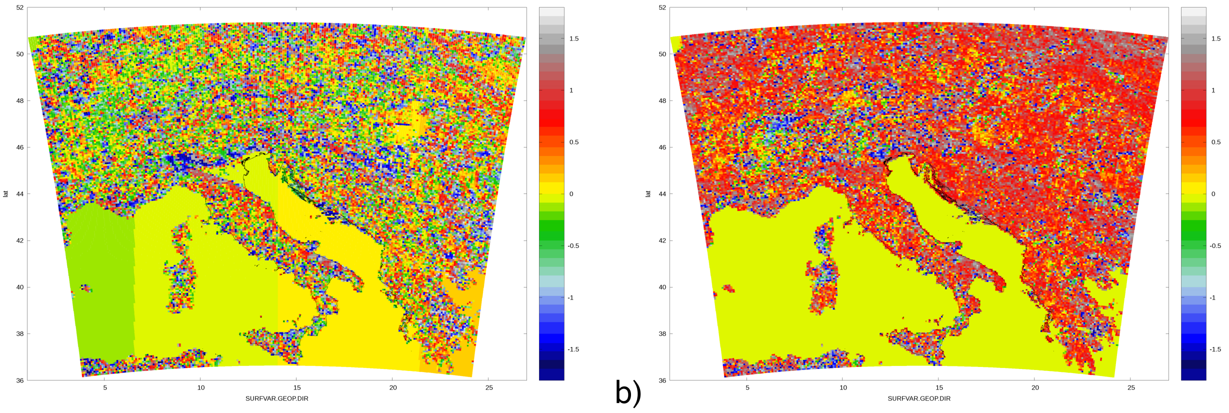
Well, this is a bit embarrassing, apparently the two angles are not defined the same way (the above derivation is valid if the absolute value of the argument of the arcus tangens function is less than unity). Let's see if it is possible to start from the SURFEX definition of theta and try to go back:

$$\theta = \frac{1}{2} atan \frac{M}{L} \qquad means \qquad \frac{M}{L} = \frac{2x}{1-x^2}$$

$$M - Mx^2 = 2Lx \qquad Mx^2 + 2Lx - M = 0$$

$$x_{1,2} = \frac{-2L + -\sqrt{4L^2+4M^2}}{2M} = \frac{-L \pm \sqrt{L^2+M^2}}{M}$$

The quadratic equation has two solutions. The one with a plus sign would correspond to the minus of the argument of the arcus tangens function got when starting from e923.



a) b)

Figure 5. The angle of the main axis of topography with the x axis, theta in ISBA (a) and in SURFEX (b) recomputed to be in radians. The values on the slopes north and south of the Po river valley seem to match the definition for ISBA with the positive values northward of x axis and negative southwards. For SURFEX it is far less clear. There is too much red colour (values between 0.5 and 1).

3.2 Step 2

This step calculates the following fields: dominant land use type, bare ground albedo (used in 3 fields), emissivity, maximum and useful depth of the soil column, percentages of clay and sand and maximum vegetation fraction.
Input datasets are global, written from 0 to 360 degrees in longitude and from north to south in latitude. The input data are prepared from different sources (Meteosat, NOAA-4, CLIMAP, ISLSCP, climatology from Webb, tables, classifications Wilson&Henderson-Sellers, Mahfouf). The names of the input files are hard-coded in the subroutine eincli2.F90. Input data files are taken from the directory:
/home/gmap/mrpa/tailleferf/923/SURFACE_G/version2/i3e

Fields are calculated with a mask, that takes into account the missing data. One mask is imposed for the bare ground albedo and emissivity in one turn and another mask for hydrological depth of soil, percentages of clay and sand. Data above sea, lake and ice-cap are assumed to be missing (undefined) for the fields such as soil depth and texture as well as for vegetation characteristics. There is no mask imposed on bare ground albedo and emissivity if there is no missing data. Information from the fraction of vegetation cover is used for root depth. The fraction of vegetation cover is weighted by the proportion of land. A 4-point interpolation operator is used for all fields.

Table 7. Files used as input for the Step 2. The aerial coverage and resolution of files used in each step of configuration e923 is listed in Table 2.

Name of file	contents
itp_GL	dominant land use type
alb_GL	Albedo of the bare ground
emi_GL	Emissivity of surface
dps_GL	Hydrological soil depth
arg_GL	Percentage of clay in the grid cell
sab_GL	Percentage of sand in the grid cell
vgx_GL	maximum vegetation cover along the annual cycle
dpr_GL	Root depth

The procedure uses land sea mask and proportion of land from the Const.Clim file created in the Step 1 of the configuration e923. This is why the output of the Step 1 was corrected using the fields from the SURFEX file before the Step 2. The output fields in the output file Const.Clim (added in the Step 2) are:

1) SURFIND.VEG.DOMI - dominant land use type, 1 for sea, 2 for ice, 3 for desert, 5 for lakes, defined by NTPMER= 1 NTPGLA= 2 NTPDES= 3 NTPLAC= 5 in YOMCLI. Default values for other fields depend on the land use type. This field is modified (overwritten) by the high resolution data in the subsequent steps of the configuration e923 (see Figure 15 a for an example).

2) SURFPROP.ARGILE - percentage of clay, varies from 3 (above the sea surface!) to 58, the field appears unnatural (Figure 6a). It can be computed from the percentage of clay from the SURFEX PGD file after correcting for the values above the sea (from 10^{20} to 0), limiting the value with the proportion of land and scaling with 100, the resulting field is shown in Figure 6b. This field could be modified by high resolution data in Step 5, but

there is no input file.

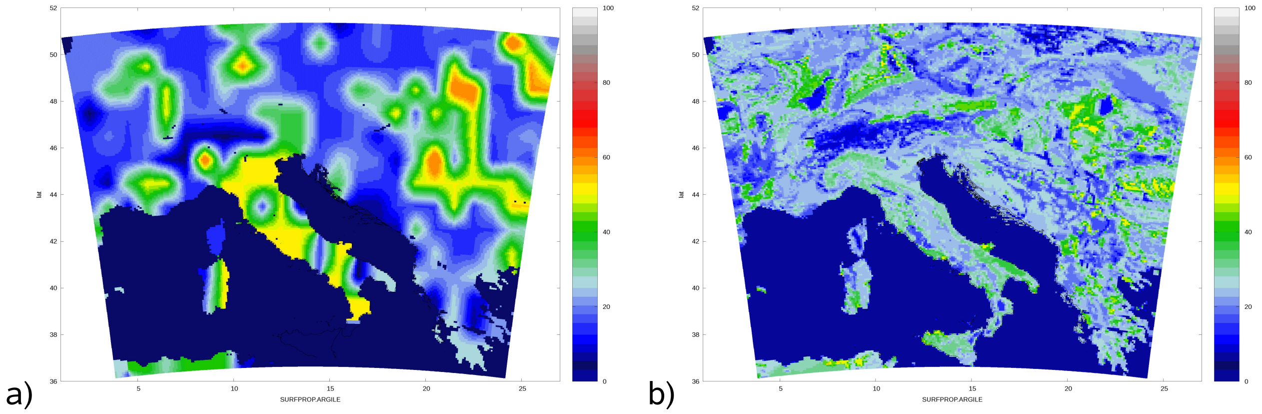


Figure 6. Proportion of clay in the clim file (a) and after the correction using data from the SURFEX PGD file (b).

3) SURFPROP.SABLE - percentage of sand, varies from 6 (above the sea surface!) to 92, the field appears unnatural (Figure 7a). It can be computed from the percentage of sand from the SURFEX PGD file after correcting for the values above the sea (from 10^{20} to 0), limiting with the proportion of land and scaling with 100, the resulting field is shown in Figure 7b. This field could be modified by high resolution data in Step 5, but there is no input file.

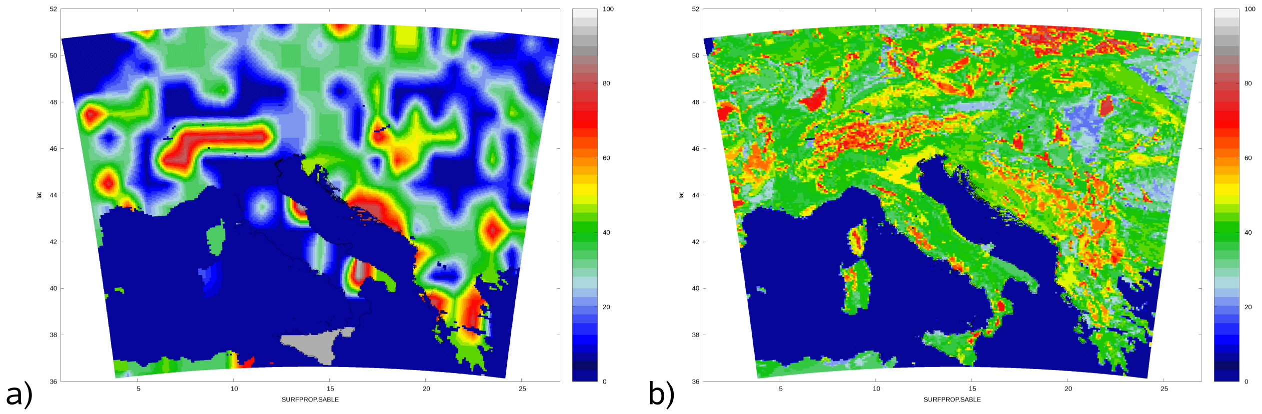


Figure 7. Proportion of sand in the clim file (a) and after the correction using data from the SURFEX PGD file (b).

4) SURFALBEDO, SURFALBEDO.COMPL, SURFALBEDO.SOLNU – albedo and bare ground albedo, all fields have the same values that exhibit rather unnatural features (see Figure 8a). The ad-hoc constructed formula is

$$A=(0.3*P_{sand}+0.1*P_{clay}+0.2*(1-P_{sand}-P_{clay}))*P_{land}+0.07*(1-P_{land})$$

and the resulting fields have the high resolution features (Figure 8b), but the coefficients in the formula above should be based on more scientific approach and possibly tuned in experiments. Finally, SURFALBEDO will be modified in subsequent steps according to the vegetation data.

a) b)

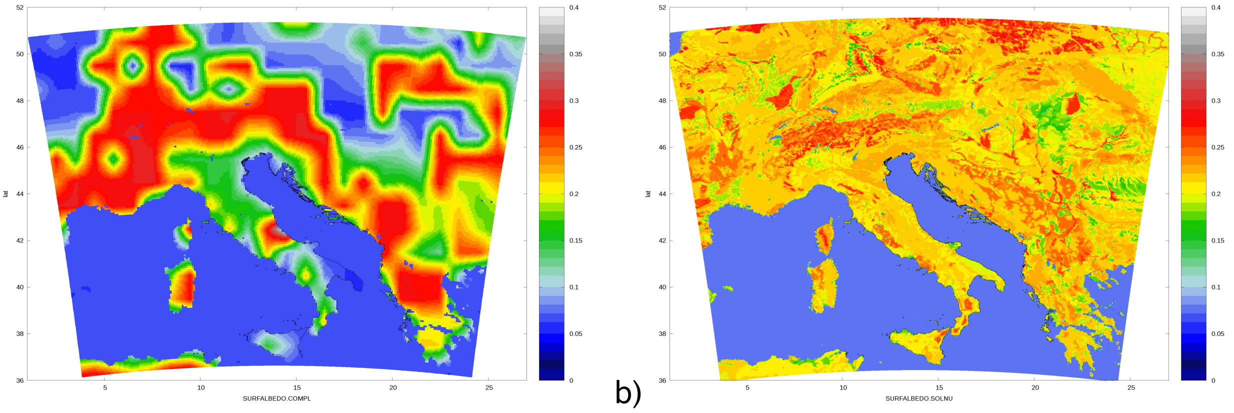


Figure 8. Albedo of bare ground in the clim file (a) and computed using the soil type data from the SURFEX PGD file (b).

3) SURFEMISSIVITE – surface emissivity, exhibits unnatural features too (see Figure 9a). This field was modified according to the land sea mask (Figure 9b) and set to one constant value above land and another above water surfaces. This field could be modified by high resolution data in Step 5, but there is no input file.

a) b)

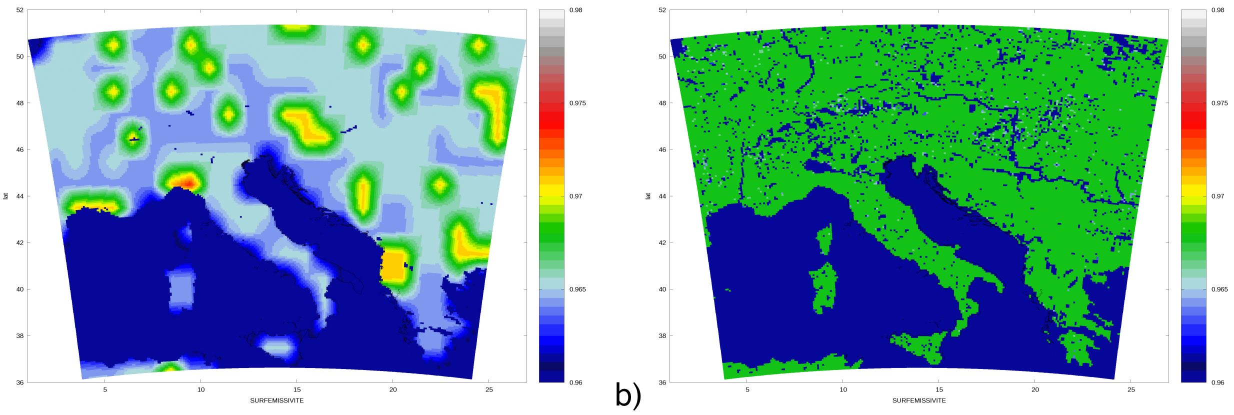


Figure 9. Surface emissivity in the clim file (a) and computed using the land sea mask data from the SURFEX PGD file (b).

4) SURFEPAI.SOL.MAX - maximum depth of the soil column that exhibited rather unnatural (and low resolution) features (Figure 10a), so this field was computed from the percentages of clay (P_{clay}) and sand (P_{sand}) and the proportion of land (P_{land}) and set to 8 for the water surfaces, the ad-hoc constructed formula is

$$EPSmax=(3*P_{sand}+5*P_{clay}+4*(1-P_{sand}-P_{clay}))*P_{land}+8*(1-P_{land})$$

and the resulting field has high resolution features (Figure 10b) that include the effect of grid cells being partially covered by water, but the coefficients in the formula above should be based on more scientific approach (for example taken from a high resolution database, or computed by a refined formula based on measurements and experiments) and possibly tuned in experiments. This field could be modified by high resolution data in Step 5, but there is no input file.

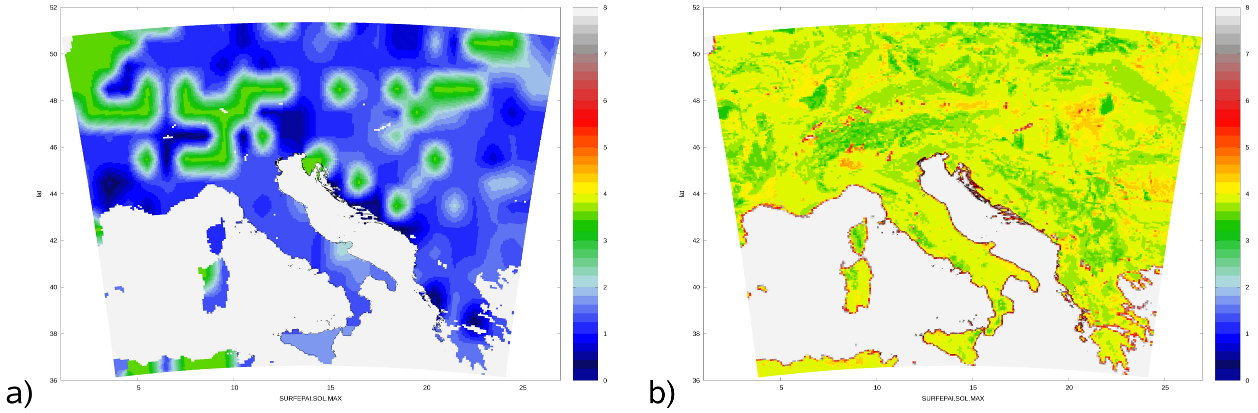


Figure 10. Maximum depth of the soil column in the clim file (a) and computed using the land sea mask data from the SURFEX PGD file (b).

9) SURFEPAIS.SOL - useful depth of the soil column is set equal to SURFEPAI.SOL.MAX in this step, so it was set to the new values of this field (described above). However, it is modified in step 5 of the configuration e923.

7) SURFPROP.VEG.MAX (maximum vegetation fraction) and SURFPROP.VEGETAT (proportion of vegetations) are set the same way in this step and modified according to the proportion of land from the previous step, but both are modified in subsequent step 5.

3.3 Step 3

In this step, for each of the 12 months, seven fields are computed: climatological surface, deep soil and relaxation temperatures and relative moisture contents as well as an empirically estimated equivalent water depth snow climatology. The depth of the deep soil layer is a model parameter that is accounted for in the routine that describes parametrization of the temperature profile in soil. The relaxation values for deep temperature and soil moisture are computed in the same way as the deep soil values. The fields of albedo, emissivity and roughness over sea are modified according to the extension of sea-ice if they are present in the initial clim file, and snow cover is modified according to land use type (ice cap or not).

The source datasets are taken from an old NCAR climatology for surface temperature and moisture, and AMIP1 mean data for sea surface temperature and sea-ice extent. Input fields are written from 0 to 360 degrees in longitude and from North to South in latitude. The target grid is a regular grid (geographical or in plane projection). A sub-grid of the final grid is introduced for interpolation. The values on the sub-grid is obtained by a 12-point interpolation operator applied on the source grid. The values on the final grid are the averages of the values in the boxes (a box is the part of the sub-grid corresponding to a point of the final grid). What is this procedure doing for the high resolution domains is a bit unclear.

The soil moisture and the amount of snow are computed from temperature, moisture and fraction of the ice cap using empirical formula. The deep soil and relaxation values are obtained by filtering the surface values in time with a factor EXP(-(1+i)H/SQRT(T)) in Fourier space, where T is the period (in days) of the wave and H the depth between the mid-surface layer and the mid-nth layer scaled by the e-folding depth of the diurnal cycle.

Input data file is copied from the directory from directory:

/home/gmap/mrpa/tailleferf/923/N108/i3e/

There is only one input data file named *N108_GL* (name also hard-coded in the source code of the subroutine eincli3.F90). The file contains the fields describing the fraction of land, height of orography multiplied by g=9.81, fraction of the land covered by ice caps, monthly fields for surface temperature, dew-point temperature and sea surface temperature. The same file contains 12 fields for the variables with a yearly cycle, such as temperature and dew-point temperature, but only one field for the ice cap fraction (glaciers).

The routine first reduces the temperature and dew-point temperature to the sea level using standard atmosphere vertical gradient of temperature and orography also from *N108_GL* file. Then land sea mask, surface geopotential, land use type, albedo, emissivity and surface roughness are read from the Const.Clim file created in the previous step.

Output fields in the 12 output files Const.Clim.$MM ($MM is number of month). Each field that existed in the input Const.Clim file is copied to each of the output files. The following fields are added or modified:

1) SURFTEMPERATURE - climatological surface temperature is computed from the temperature reduced to the mean sea level by using the standard atmosphere vertical gradient of temperature and topography from the Const.Clim file (Figure 11a).

2) PROFTEMPERATURE - climatological deep soil temperature, the depth of soil is controlled by the ZTSHT parameter, that is in turn computed from the SODELX values in YOMPHY1, the values are computed by the routine ***inclag*** and the values above the sea are set equal to the surface values (Figure 11b).

3) RELATEMPERATURE - relaxation values for deep temperature are set equal to the deep soil values (Figure 11b).

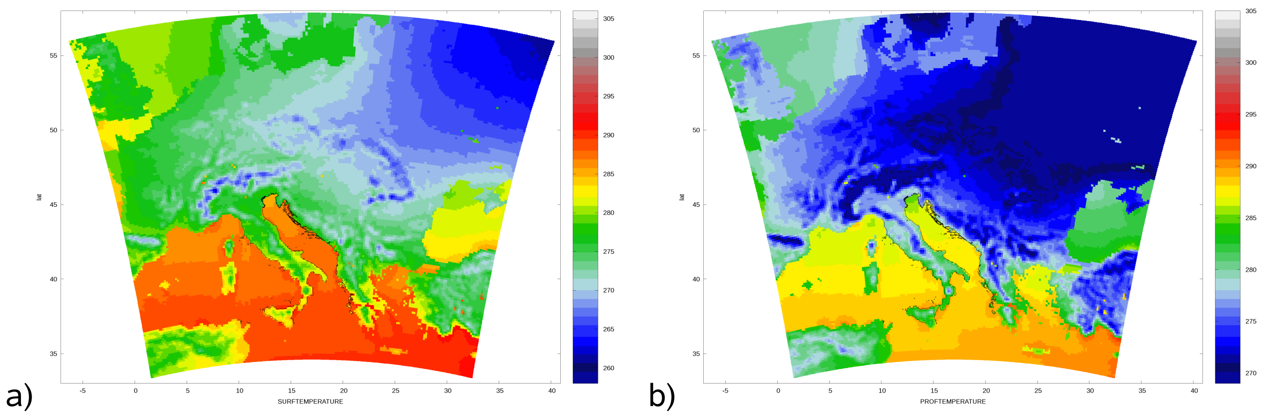


Figure 11. Surface (a) and soil (b) temperatures after the Step 3 for January.

4) SURFPROP.RMAX.EA - climatological surface relative moisture content is computed in the following way. First the input dew-point temperature is reduced to the mean sea level using the standard atmosphere vertical gradient of temperature and topography from the N108_GL file. Then the climatological surface dew-point temperature is computed by raising it using the standard atmosphere vertical gradient of temperature and topography from the Const.Clim file. Then relative humidity is then computed using the function

$$RH=e_s(T_d)/e_s(T)$$

where e_s is the saturation pressure for the water vapour, T and T_d temperature and dewpoint temperature. Then the soil relative water content is computed from the relative humidity using the function

FOWR=acos(1-2RH)/pi

and these computations are repeated for each monthly field (Figure 12a).

5) PROFPROP.RMAX.EA - climatological deep soil relative moisture content, the depth of soil is controlled by the PPWSHT parameter, the value is hard-coded to 0.1969 in eincli3.F90, the values are computed by the routine *inclag* and the values above the sea are set equal to the surface values (Figure 12b).

6) RELAPROP.RMAX.EA - relaxation values for deep soil moisture are set equal to the deep soil values (Figure 12b).

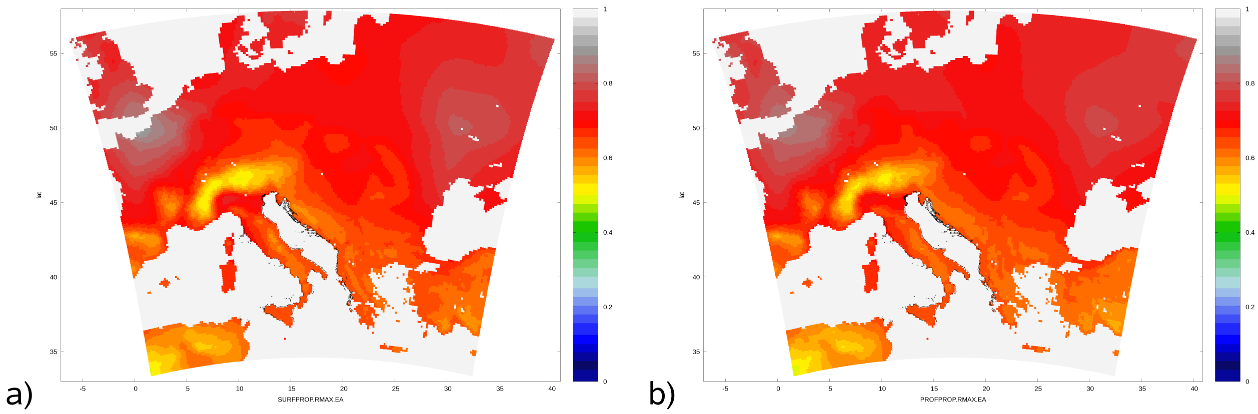


Figure 12. Surface (a) and soil (b) relative moisture contents after the Step 3 for January.

7) SURFRESERV.NEIGE - an empirically estimated climatology of the equivalent water depth for snow is computed using the ice cap fraction (ICF) from the *N108_GL* file ad the surface relative moisture content with the formula

$$SN=10*ICF/(1-ICF)+1000*(1-\cos(pi*0.04*FOWR*(T_t-T_s))$$

where T_t is the triple point temperature, T_s surface temperature, ICF is the ice cap fraction and if $T_s>T_t$, the argument of the cosinus term is set to zero, so in the case the surface temperature is larger than T_t the snow amount is controlled by the ice cap fraction (Figure 13).

8) SURFEMISSIVITE - emissivity (monthly) modified over sea only according to the extent of the sea ice, if there is no sea ice over the domain (in climatological average) then no effective modification (values set to default for sea/lake).

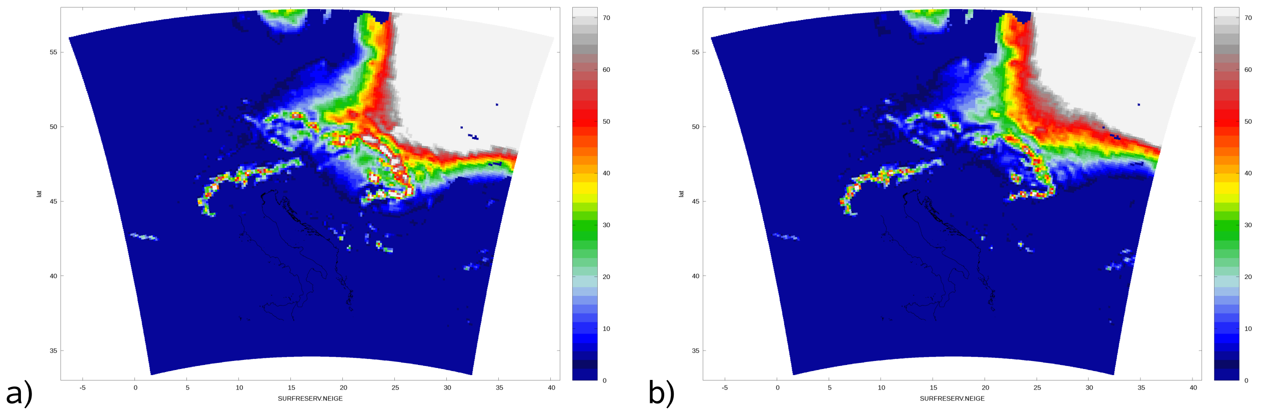


Figure 13. Snow content for January (a) and February (b).

9) SURFALBEDO - albedo (monthly) modified over sea only according to the extent of the sea ice, if there is no sea ice over the domain (in climatological average) then no effective modification (values set to default for sea/lake).

10) SURFZ0.FOIS.G – surface roughness (monthly) modified over sea only according to the extent of the sea ice, for all points where the proportion of land is less that 0.5, the values are set to the default values for sea. This might be detrimental for areas with unresolved islands, especially if these islands carry meteorological stations useful for verification and data assimilation.

3.4 Step 4

In this step, routine *eincli4* calculates monthly climatological fields for fraction of vegetation cover, kinetic roughness length of vegetation times g=9.81, vegetation albedo, leaf area index, minimum surface resistance and thermal roughness length, the values for albedo and roughness length are modified with the vegetation values.

Input data are global, organized from 0 to 360 degrees in longitude and from North pole to South pole in latitude. The subroutine uses a number of input files listed in Table 8, and the Const.Clim file from the previous step. This step is run once for each monthly file and the output file from the previous step for the corresponding month has to be copied into the Const.Clim file. The output is also written to the file with the same name. Input datasets are prepared from Henderson-Sellers classification and tables. Input data are taken from directory:
/home/gmap/mrpa/tailleferf/923/SURFACE_G/version2/i3e

The fields of maximum vegetation fraction, bare ground albedo, roughness length of topography and land use type are used from the Const.Clim file from the previous step (Const.Clim.$MM, where MM is the two digit number of the month). The input files for leaf area index and fraction of vegetation are named lai$MM_GL and veg$MM_GL in the input directory and the file for the corresponding month has to be copied into lai_GL and veg_GL respectively. Be aware that the input files for Step2 and Step4 are stored in the same directory.

Table 8. Files used as input for the Step 4, number of files for the whole year and their contents. The aerial coverage and resolution of files used in each step of configuration e923 is listed in Table 2.

File name	Files	Contents

veg_GL	12	fraction of vegetation cover
z0v_GL	1	roughness length of vegetation
alv_GL	1	albedo of vegetation
lai_GL	12	leaf area index
rsm_GL	1	minimum surface resistance

Fields are computed with a mask, taking into account missing data. A 4-point interpolation operator is used for all fields. Instead of interpolating surface roughness and minimum surface resistance, the values taken into the interpolation routine are logarithm of the surface roughness and the ratio of the leaf area index and minimum surface resistance. The fraction of vegetation cover is weighted by the fraction of land.

Vegetation characteristics are taken into account only as long as the fraction of vegetation cover keeps significant (meaning that the fraction of vegetation is larger than SVEG = 0.02 that is a tuning parameter that can be set in NAMCLI). The ratio of thermal to kinetic roughness length is fixed to 1/10 (if LZ0THER=T). This is also important since it is related to the choices that have to be made in the subsequent 001 *namelist*.

The following output fields are added or modified in the output file Const.Clim file:

1) SURFPROP.VEGETAT - fraction of vegetation cover is read from the input file and interpolated with mask, while the values where the fraction of vegetation is less than 0.02 are defined as desert. The proportion of vegetation is finally corrected using the proportion of land:

$$P_{veg}=\min(P_{veg},P_{veg\_max}*P_{land})$$

where P_{veg} is the proportion of vegetation (read on the input file), P_{veg_max} is the maximum proportion of vegetation (read from the Const.Clim file) and P_{land} is the proportion of land in the grid cell.

2) SURFZ0VEG.FOIS.G – is kinetic roughness length of vegetation multiplied by g=9.81m/s2. The vegetation roughness length (Z_{0veg}) is read from the input file, interpolated using the missing data mask followed by the desert and the land sea mask. The logarithm of z_0 is interpolated. Then the field is modified using the proportion of vegetation P_{veg} and proportion of urban areas P_{urb}:

$$z_{0veg}=SFCZ0*g*(P_{veg}*z_{0veg}+P_{urb}*2.5+(1-P_{urb}-P_{veg})*0.001)$$

where 2.5 and 0.001 are values set by SZZ0U and SZZ0D that are used from YOMCLI and the z_{0veg} can be scaled by SFCZ0 that is set to 1. Consequently, *vegetation roughness length also includes contribution from the urban areas*.

3) SURFALBEDO.VEG – albedo of vegetation is read from the input file, interpolated using a missing data mask and additionally with the land sea mask and desert mask.

4) SURFIND.FOLIAIRE - leaf area index is read from the input file and then interpolated using a mask, the possible negative values are put to be zero.

5) SURFRESI.STO.MIN - minimum surface resistance is read from the input file, the ratio of leaf area index and minimum surface resistance is interpolated.

6) SURFZ0.FOIS.G - global roughness length is changed the contribution from vegetation and urban areas are added through the formula:

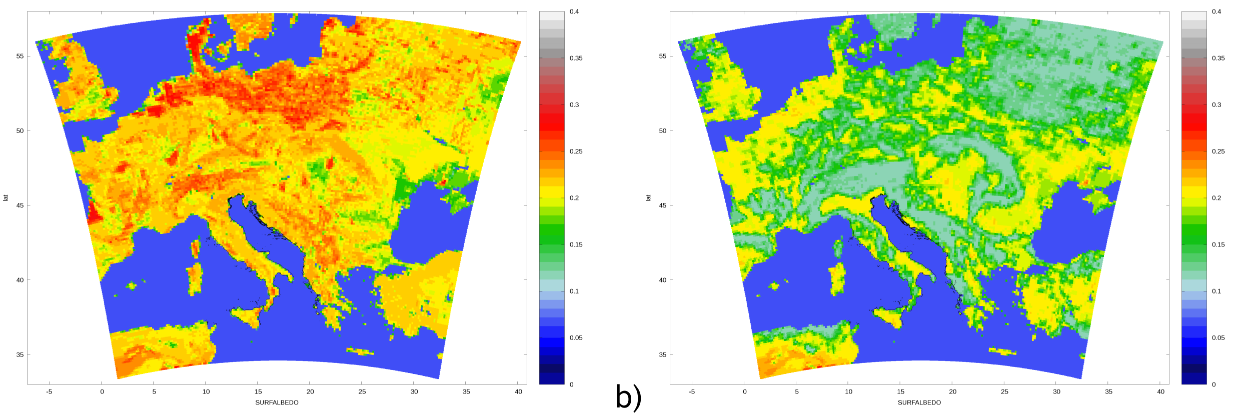
$$z_0=sqrt(z_0^2+z_{0veg}^2)$$

7) SURFGZ0.THERM - thermal roughness length multiplied by g=9.81m/s2 is in the case of LZ0THER=T equal to STHER*z_0, where STHER=0.1 is a tuning parameter, and if LZ0THER=F, then it is equal to STHER*z_{0veg}. However, what is the value for STHER in this case? Preliminary tests show that STHER=1 yields better scores.

8) SURFALBEDO – albedo A that was equal to the bare ground albedo A_{bg} until this step is modified using the vegetation albedo A_{veg} and proportion of vegetation P_{veg}

$$A=(1-P_{veg})*A_{bg}+P_{veg}*A_{veg}$$

The fields are changed to default values for desert if P_{veg}<SVEG=0.02, default values for sea and glaciers if the main land use type is sea, lake or glacier (Figure 14).



a) b)

Figure 14. Input surface albedo (a) and output (b) when the vegetation albedo is included into the total albedo.

3.5 Step 5

This step modifies fields that already exist in the Const.Clim file. The climatological fields describing land surface are modified in this step using the high resolution input data available from a limited rectangular domain. Fields computed in Step2 and Step4 can be changed if there is corresponding high resolution input data available. Fields computed in Step1 (topography) and Step 3 (temperature) are not changed in this step. Up to 17 fields which characterize vegetation and soil can be modified: percentage of vegetation, maximum percentage of vegetation, maximum soil depth, land use type, leaf area index, minimum surface resistance of vegetation, useful soil depth, bare ground albedo, albedo of vegetation, albedo, emissivity, thermal and kinetic roughness lengths, percentages of clay and sand.

The target grid is a regular grid (geographical or in plane projection). The source grid can be any regular rectangular latitude by longitude grid, the first point being at the north-west edge, longitudes going eastwards, and latitude going southwards (the north-east edge is before the south-west edge). The earth is supposed to be flat, i.e. the periodicity of the longitudes as well as the symmetry of the latitudes is ignored. It can be used for a global source grid (although it has not been designed for this task), but the interpolation at the boundaries will not be optimal. Inside the input domain, the variables are averaged in a rectangular latitude longitude box around each point of the target grid; the size of the box is approximately twice the distance between two points in the target grid. The missing data have a negative value, and it is assumed that the location of missing data is the same for each field. Over parts of the output domain where the input data are missing, the original value is kept. At the boundaries of the input data domain, a linear combination of the old and new values is performed. Values are not modified if the corresponding input file is not available.

Table 9. Files used as input for the Step 5, and their contents. The first column contains the name of the input file and the availability of the file (NO if it is not available, 1 for one file for whole year and 12 for separate file for each month). The area coverage and resolution of files used in each step of configuration e923 is listed in Table 2.

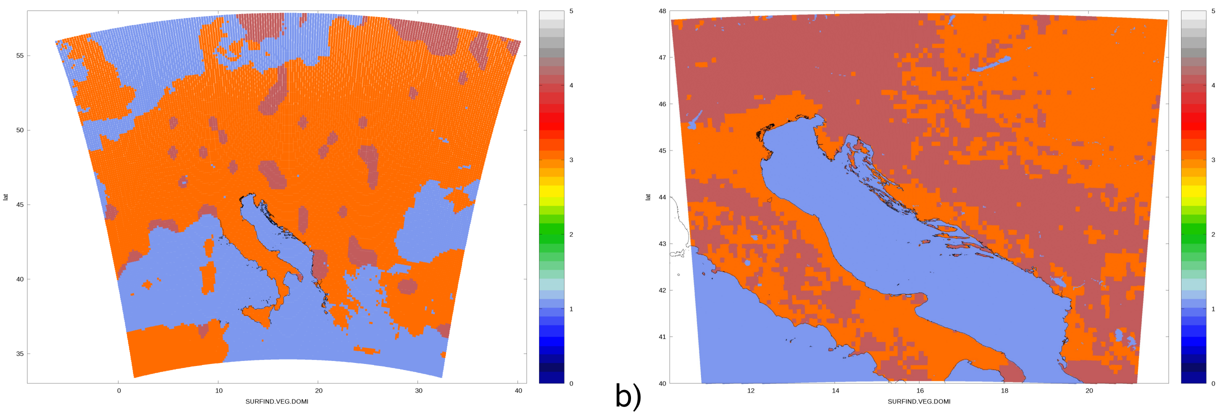
File name	contents
msk_HR / 1	mask describing missing data, -9999. -> missing 1. -> available
als_HR / NO	bare ground albedo (no input file)
emi_HR / NO	emissivity
arg_HR / NO	percentage of clay
sab_HR / NO	percentage of sand
dps_HR / NO	maximum soil depth
itp_HR / 1	dominant land use type
vgx_HR / 1	maximum vegetation cover
veg_HR / 12	current vegetation cover
lai_HR /12	leaf area index
rsm_HR / 1	miminum surface resistance
z0v_HR / 1	roughness length of vegetation
alv_HR / 1	albedo of vegetation
dpr_HR / 1	root depth

Input datasets are result from the combination of a AVHRR-derived vegetation mapping over Europe and the ESA forest mask. The initial data have a resolution of 2 km but have been interpolated onto a regular 0.1x0.1 degrees grid. The area is ELATSW=30, ELONSW=-25, ELATNE=72, ELONNE=61. Input data files are copied from the directory:
/home/gmap/mrpa/tailleferf/923/SURFACE_L/EUROPEb_v1/i3e

The routine first checks for the existence of each of the xxx_HR input files. It stops execution only if there is no file with a common mask. Further conditions should be satisfied:
- both high resolution sand and clay should be available (otherwise none of them is used), however, none of them is available,
- both high resolution maximum and current vegetation cover should be available (otherwise none of them is used),
- both high resolution leaf area index and minimum surface resistance should be available (otherwise none of them is used).

The following two warnings are issued (in the NODE file and err output) because there is no input high resolution data on emissivity:
CAUTION : IF VEGETATION FRACTION IS MODIFIED BY EINCLI5, ALL VEGETATION CHARACTERISTICS SHOULD BE MODIFIED !
CAUTION : IF LAND USE TYPE IS MODIFIED BY EINCLI5, ALL VEGETATION CHARACTERISTICS SHOULD BE MODIFIED !
and the next warning is issued because there are no high resolution files for sand and clay
CAUTION : SOIL DEPTH MODIFIED BY EINCLI5 BUT NOT SOIL TEXTURE .

The Const.Clim (Const.Clim.$MM where MM is the corresponding month) from previous step is used as a source of data for all the fields computed in this step. The following output fields in the output file Const.Clim are modified if there is available input data in high resolution:
SURFIND.VEG.DOMI - land use type
SURFPROP.VEGETAT - percentage of vegetation
SURFPROP.VEG.MAX - max. percentage of vegetation
SURFIND.FOLIAIRE - leaf area index
SURFRESI.STO.MIN - minimum surface resistance of vegetation
SURFEPAIS.SOL - useful soil depth
SURFZ0.FOIS.G - global roughness length
SURFZ0VEG.FOIS.G - Vegetation roughness length (*g)
SURFGZ0.THERM - thermal roughness length
SURFALBEDO – albedo (combined bare land and vegetation)
SURFALBEDO.SOLNU - Albedo of bare ground (no input data)
SURFALBEDO.COMPL – Surf ace albedo for non snowed areas.
SURFEMISSIVITE – emissivity (no input data)
SURFPROP.ARGILE - percentage of clay (no input data)
SURFPROP.SABLE - percentage of sand (no input data)
SURFEPAI.SOL.MAX - maximum soil depth (no input data)
SURFALBEDO.VEG - vegetation albedo
SURFPROP.URBANIS - Proportion of urbanisation



a) b)

Figure 15. Land use type for the LACE coupling domain after Step 2 (a) and on a smaller domain in 2 km resolution (b) after Step 5.

First the mask file is read and then all the other files that exist. Then the Const.Clim file is read for all of the variables that can be modified due to available data. The final fields are computed as a linear combination using the weights w1 for the high resolution input data and w2 for the data from the Const.Clim file. If there are no new (high resolution) fields, the old values are kept. The new value is taken (weight is one) from the high resolution dataset if all points in the interpolation box have non-missing values. If all points are missing then the weight for the value from the Const.Clim file is one. When the choice of the dominant type of vegetation is computed, the condition is that the dominant type in the high resolution dataset exceeds 25%, otherwise the vegetation type from the Const.Clim file is kept.

3.6 Step 6

This step modifies fields that already exist in the Const.Clim file (created in Step 3). Fields computed in Step 1 (topography), 2 and 5 are not changed. The climatological fields are modified in this step using the input data available on a global domain resulting from an ARPEGE experiment. The following fields are recomputed using an alternative dataset: climatological mean soil, sea and land surface temperature, surface and soil moisture and an empirically estimated equivalent water depth snow climatology. Runs with the following geographical settings of the input data set in the *namelist*:

```
 &NAMCLI
   LZ0THER=.FALSE.,
   NDATX=360,
   NDATY=180,
   RSTR=0.,
```

Input data files are copied from the directory (see table 10 for details)
/home/gmap/mrpa/tailleferf/923/CLIM_G/version2/i3e

One should notice that **the same input file is used for the surface and soil temperature,** and another **single file is used for the surface and soil moisture**. Therefore, the resulting fields for all temperatures (surface SURF, soil PROF and relaxation RELA) are the same, the same is valid for the moisture fields. The usage of the sea surface temperature input file requires the existence of the land sea mask file (none of them is used). The usage of surface and soil temperatures requires the existence of the input file for topography.

Table 10. Input files for step 6. All files but topography have 12 versions for the 12 months in a year.

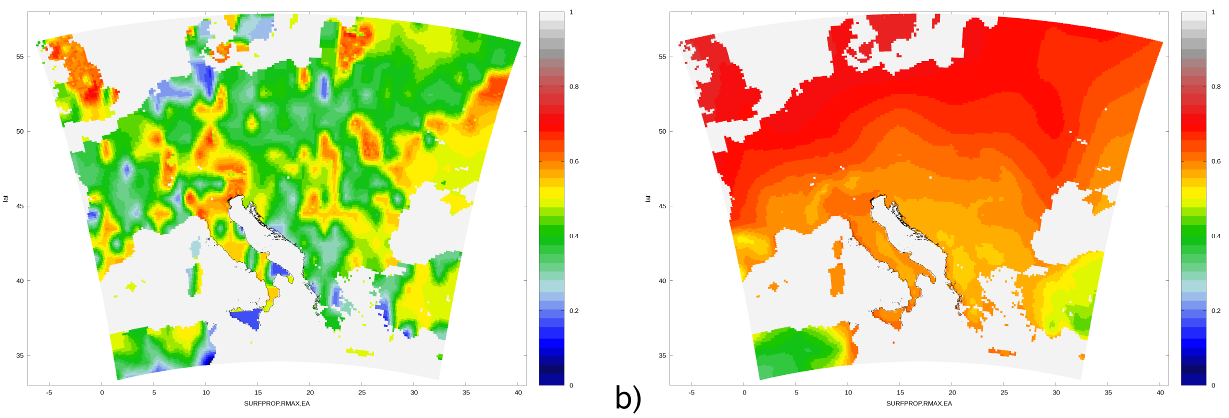
Filename	Contents (and file that is actually used)
snl_GL	Snow cover (snl_$MM_GL)
tpl_GL	Mean surface temperature over land (tpl_$MM_GL)
tsl_GL	Land surface temperature (tpl_$MM_GL)
wpl_GL	Mean soil moisture (wpl_$MM_GL)
wsl_GL	Surface soil moisture (wpl_$MM_GL)
ltm_GL	Land sea mask (no input file)
rel_GL	Topography height multiplied by g=9.81m/s2 (rel_GL, one file)
sst_GL	Sea surface temperature (no input file)

The routine prints into the node file which fields are to be modified. It is possible to use only surface temperature file and not the soil temperature file, on the other hand if there is only the soil temperature file, the surface temperature is set equal to the soil temperature.

The routine first checks for the existence of the input files. The input files are read (those that do exist). Temperatures from the input files are first reduced to the mean sea level using standard atmosphere vertical profile for temeprature. Then, the temperature fields are interpolated to the model grid of the Const.Clim file and then the temperature is raised to the terrain height using the same vertical profile. If there is input sea surface temperature, values for albedo, emissivity and roughness lengths are set to default values for sea ice if the sea surface temperature is below 273.16 and to default values for sea surface otherwise (however, the point for freezing of the salty sea is about -2C so this could be refined). If temperature (surface or soil) is larger than 273.16 or soil moisture is less than RSWR (a tuning parameter set to 0.02 by default), then there is no snow. Since there is no data file for the sea surface temperature, albedo, emissivity and roughness lengths are left unchanged.

The Const.Clim (Const.Clim.$MM where MM is the corresponding month) from the previous step is used as a source of data for the fields computed in this step. The following fields are read from the Const.Clim file: vegetation index, proportions of sand and clay, useful soil depth, albedo, emissivity, kinetic and thermal roughness lengths, surface and soil temperatures and moistures and snow amount. These output fields in the output file Const.Clim are modified in the case of available input data:

SURFTEMPERATURE - climatological sea/land surface temperature
PROFTEMPERATURE - climatological mean soil temperature
RELATEMPERATURE - relaxation values for deep temperature
SURFPROP.RMAX.EA - climatological surface moisture
PROFPROP.RMAX.EA - climatological soil moisture
RELAPROP.RMAX.EA - relaxation values for deep soil moisture
SURFRESERV.NEIGE - an empirically estimated equivalent water depth snow climatology



a) b)

Figure 14. Climatological surface moisture after the step 6 (a) and after the Step 5 (b) for August.

3.6.1 Correction of the fields after the Step 6

Finally, the output file of the Step 6 contains three fields describing temperature (but all three hold the same values) and three fields for moisture (also the same), while the input Const.Clim file (from Step 5) also contained the three fields, but relaxation (prefix RELA) and soil (prefix PROF) were different from the surface (prefix SURF) values. Additionally the moisture fields exhibit a pattern that does not look very physical (Figure 14a), possibly due to the resolution of the input data. After the Step 6, amount of snow (Figure 15a) is rather different than after the Step 5, possibly due to low resolution of the input data (where these mountains would also be low and therefore warmer and therefore with little snow cover), on the other hand, the snow cover in the Const.Clim file before the Step 6 was an estimate using an empirical function that used the surface temperature that was lower since the mountains were higher (in higher resolution).

Therefore, as an addition, the following corrections are made after the computations of the Step 6. The difference between the soil and surface temperature in the output file of the Step 5 is computed for each monthly file and then added to the soil temperature in the corresponding output file of the Step 6. The same is done for the relaxation temperature.

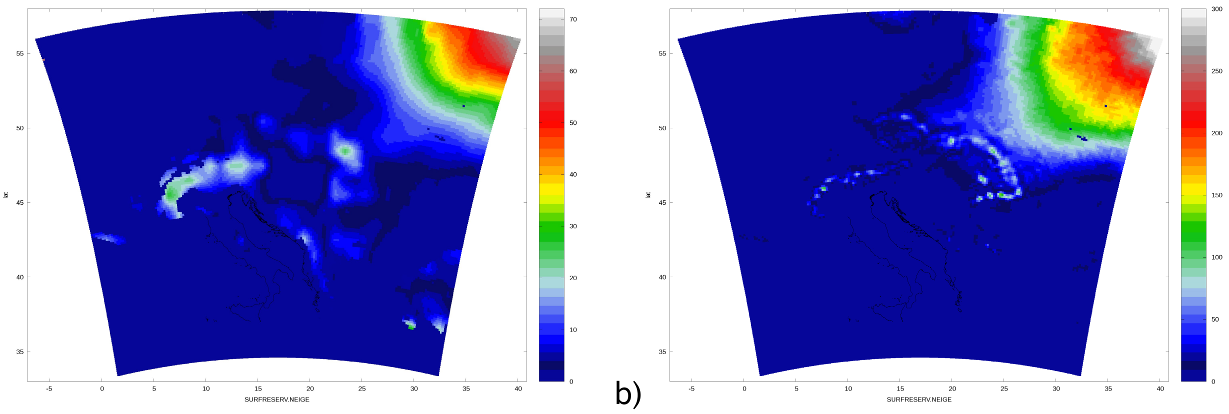
$$T_{p\_6}=T_{s\_6}+T_{p\_5}-T_{s\_5}$$

where T_{p_5} and T_{s_5} are the soil and surface temperatures output from the Step 5, T_{p_6} and T_{s_6} are the soil and surface temperatures output from the Step 6.

The fields describing surface, soil and relaxation moistures are simply taken from the output file of the Step 5 (until someone gets a better idea). The amount of snow (SN) is corrected so that the final value is the average of the empirical estimate and the data from the Step 6:

SN_6=0.5(SN_5+SN_6)

where SN_5 and SN_6 are the amounts of snow in the output files of the Step 5 and Step 6 (Figure 14b). The procedure is repeated for each monthly file and the result for January is show in Figure 15.



a) b)

Figure 15. Amount of snow after the step 6 (a) and after the correction (b) for January.

3.7 Step 8

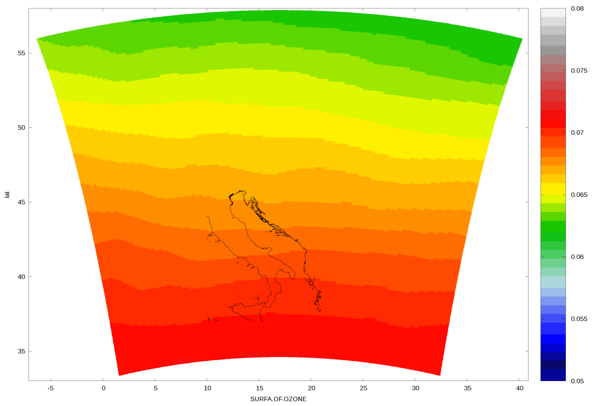
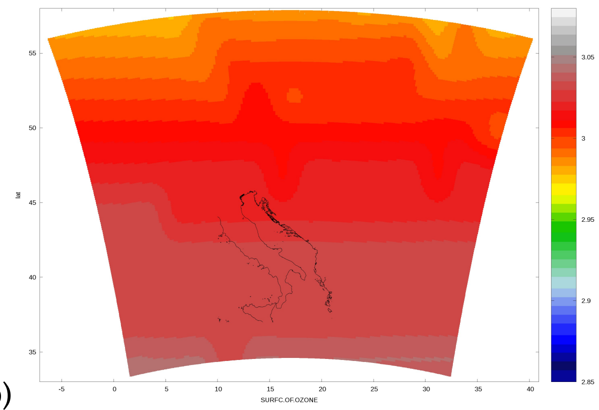
Three new fields, describing the vertical profiles of ozone, are added to the clim file in this step. Input data files are copied from the directory
/home/gmap/mrpa/tailleferf/923/CLIM_G/ozone/ascii/

The input file should be named abc_coef, it is in fact a copy of abc_quadra_${MM} from the above directory. The Const.Clim output from the previous step is also used (Const.Clim.$MM) for each month.

Table 11. New fields in the output file Const.Clim, monthly variability and values of minima and maxima.

Name of field	**Contents**	**variability**	**Values**
SURFA.OF.OZONE	First ozone profile (A)	monthly	0.054 – 0.071
SURFB.OF.OZONE	Second ozone profile (B)	constant	3166
SURFC.OF.OZONE	Third ozone profile (C)	monthly	2.85 – 3.075

Since the input fields are in rather low resolution, there is little variability in these fields (Figure 16), but there is an annual cycle for A and C.


a)

b)
Figure 16. Fields for the A (a) and C (b) coefficients of the ozone profile for April.

3.8 Step 9

Four new fields, describing four kinds of aerosols are added to the clim file in this step. Input data files are copied from the directory
/home/gmap/mrpa/tailleferf/923/CLIM_G/aerosols/ascii/

The input file should be named aero_GL, it is in fact a copy of aero.tegen.${MM} from the above directory. The Const.Clim output from the previous step is also used (Const.Clim.$MM) for each month.

Table 12. New fields in the output file Const.Clim, all fields have annual cycle and values of minima and maxima.

Name of field	Contents	Values
SURFAEROS.SEA	Marine aerosols	0.001-0.01
SURFAEROS.LAND	Continental aerosols	0.0279-0.274
SURFAEROS.SOOT	Carbon aerosols	0.0025-0.042
SURFAEROS.DESERT	Desert aerosols	0.0024-0.18

The resulting fields have rather low spatial variability (due to low resolution input data) but one should keep in mind that these fields represent a monthly value for a climatological background, and as such they might be as smooth as they currently are.

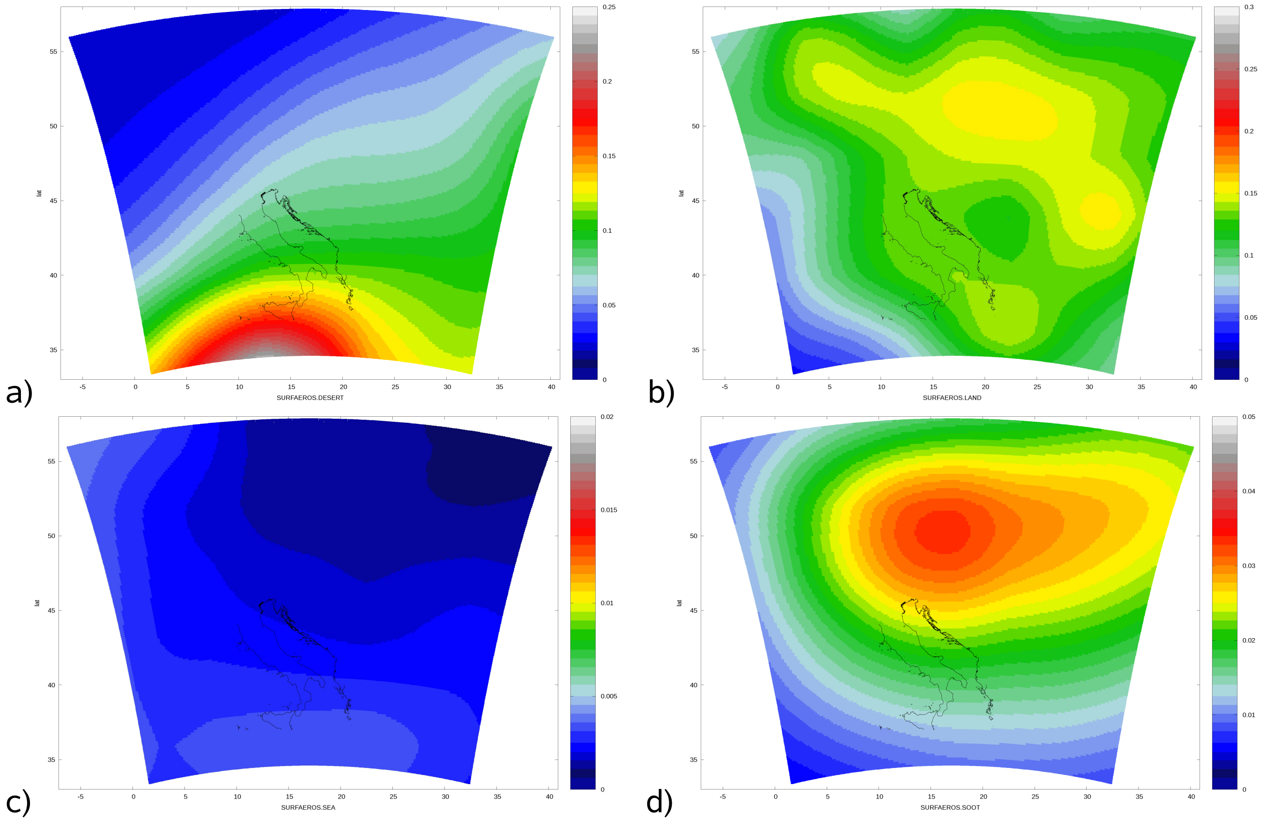


Figure 17. Desert (a), land (b), sea (c) and soot (d) aerosols for May.

Appendix A

Table A. List of important *namelist* parameters:

Namelist	Parameter	Value	meaning
NAERAD	LSRTM	F	No RRTM
NAMCT0	NCONF	923	
	CNMEXP	'CLIM'	
	LECMWF	F	
	LELAM	T	
	LRPLANE	T	
NAMDIM	NDGLG	216	Size of domain in the y direction
	NDGUXG	205	Size of the c+i zone in the y direction
	NDLON	240	Size of domain in the x direction
	NDLUXG	229	Size of the c+i zone in the x direction
	NFLEVG	1	number of model levels in grid point space
	NMSMAX	79	truncation order in longitude
	NSMAX	71	truncation order in latitude
NAMGEM	NHTYP	0	Type of grid, 0 for regular grid
	NSTTYP	1	Pole of stretching
	RMUCEN	1	Mu of the pole of stretching
	RLOCEN	0	Longitude of the pole of stretching
	RSTRET	1	Stretching factor
NAMMCC	N923	1/2/3/4/5/6/ 8/9	Step of e923
NAMPAR0	NPROC	1	**Runs only on one processor!**
NAMPAR1	LSPLIT	.FALSE.	
NAMPHY	LMPHYS	T	Global switch for MF (or non EC?) physics
	LSOLV	T	Noilhan Planton soil and vegetation scheme
	LVGSN	T	Combine vegetation and snow
NAMPHY0	GCVPSI	0	Adimensional coefficient to go from local turbulent fluxes (1) to integral (0)
NEMDIM	NBZONG	8	half-difference between the size of inner and outer zone in x direction (I zone width)
	NBZONL	8	half-difference between the size of inner and outer zone in y direction
NEMGEO	ELON0	14.000	geographic longitude of reference for the projection
	ELAT0	44.000	geographic latitude of reference for the projection
	ELONC	14.000	geographic longitude of the centre of domain
	ELATC	44.000	geographic latitude of the centre of domain
	EDELX	8000	grid size in meters along x
	EDELY	8000	grid size in meters along x

Appendix B

Table B. Fields in climate files and their properties (in the clim files until 2015).

Variable name	Description	year	sea	Range	Quality and correction
SURFIND.TERREMER	Land/sea mask	no	0	0 - 1	Extra land points
SURFGEOPOTENTIEL	OUTPUT gp orography (*g)	no	-	-192- 29775	ok
SURFET.GEOPOTENT	Std. dev. of orography (*g)	no	0	0 - 9246	ok
SURFVAR.GEOP.ANI	Anisotropy of topography	no	1	0 - 1	ok
SURFVAR.GEOP.DIR	Topography main direction	no	-	-1.71–1.71	ok
SURFZ0.FOIS.G	Surface roughness (*g)	yes	0.001	0.001-114.7	10*gz0therm, smooth
SURFZ0REL.FOIS.G	Roughness length (*g)	no	0.001	0.001-113.8	Chessboard, smooth
SURFGZ0.THERM	Heat roughness length (*g)	yes	0.001	0.001-11.47	Chessboard, smooth
SURFZ0VEG.FOIS.G	Vegetation roughness length	yes	0	0-18.87	Chessboard, smooth
SURFTEMPERATURE	Surface temperature (K)	yes	-	259-305	ok
PROFTEMPERATURE	Deep soil temperature	yes	-	259-305	Values as surftemp.
RELATEMPERATURE		Yes		259-305	Values as surftemp.
RELAPROP.RMAX.EA	Relax. deep soil wetness	yes	1	0.14-1	Could use improvement
SURFPROP.RMAX.EA	Clim. rel. surf soil wetn	Yes	1	0.14-1	values=relaprop.rmax.ea
PROFPROP.RMAX.EA	Clim. rel. deep soil wetn	Yes	1	0.14-1	values=relaprop.rmax.ea
SURFRESERV.NEIGE	Snow depth	yes		0-71	Could use improvement
SURFALBEDO	Albedo	yes	0.07	0.07-0.35	
SURFEMISSIVITE	Emissivity	no	0.96	0.96-0.975	Could use improvement
SURFALBEDO.COMPL	Surf. albedo for non snow	no	0.07	0.05-0.38	vals = surfalbedo.solnu
SURFALBEDO.SOLNU	Surf. albedo for bare ground	no	0.07	0.05-0.38	vals = surfalbedo.compl
SURFALBEDO.VEG	Surf. albedo for vegetation	no	0.05	0.05-0.2	ok
SURFPROP.VEG.MAX	Maximum proportion of veg	no	0	0-0.91	
SURFPROP.VEGETAT	percentage of vegetation	yes	0	0-0.91	Can > surfprop.veg.max
SURFPROP.URBANIS	Proportion of urbanisation	no	0 (-)	0-0.383	ok
SURFPROP.TERRE	percentage of land	no	<0.5	0-1	Correct for peninsulas
SURFPROP.ARGILE	Percentage of clay in soil	no	3	3-58	Low res, >0 over sea
SURFPROP.SABLE	Percentage of sand in soil	no	6	6-92	Low res, >0 over sea
SURFEPAIS.SOL	Soil depth	no	8	0.7-8	Larger than maximum
SURFEPAI.SOL.MAX	Maximum soil depth	no	8	0.1-8	Low resolution
SURFIND.VEG.DOMI	Index of vegetation	no	1	1-4	
SURFRESI.STO.MIN	Stomatal minimum resistance	yes	5000	40-200	
SURFIND.FOLIAIRE	Leaf area index	yes	0	0.14-4	
SURFAEROS.SEA	Marine aerosols	yes	-	0.001-0.01	Low resolution
SURFAEROS.LAND	Continental aerosols	yes	-	0.027-0.275	Low resolution
SURFAEROS.SOOT	Carbon aerosols	yes	-	0.002-0.045	Low resolution
SURFAEROS.DESERT	Desert aerosols	yes	-	0.002-0.18	Low resolution
SURFA.OF.OZONE	First ozone profile (A)	yes	-	0.054-0.072	Low resolution
SURFB.OF.OZONE	Second ozone profile (B)	no	-	3166	Always = 3166.000000
SURFC.OF.OZONE	Third ozone profile (C)	yes	-	2.85-3.08	Low resolution