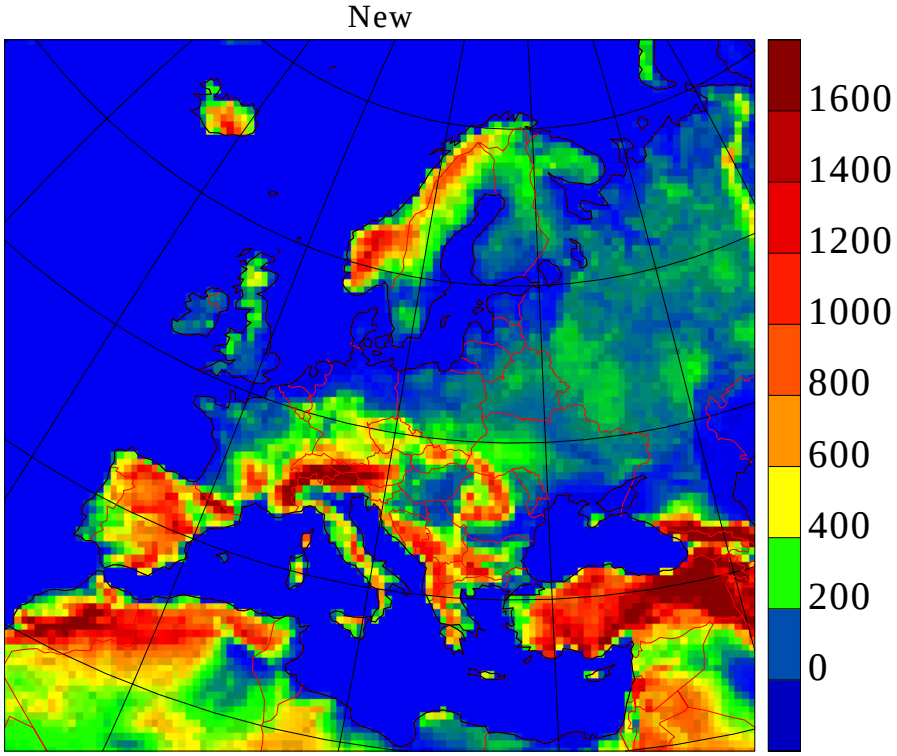
DANISH METEOROLOGICAL INSTITUTE

——— TECHNICAL REPORT ———

01-15

High resolution physiographic data set for HIRHAM4: An application to a 50 km horizontal resolution domain covering Europe

Jens Hesselbjerg Christensen, Ole Bøssing Christensen, Jan-Peter Schulz, Stefan Hagemann, Michael Botzet



COPENHAGEN 2001

ISSN: 0906-897X (printed version)
1399-1388 (online version)

High resolution physiographic data set for HIRHAM4: An application to a 50 km horizontal resolution domain covering Europe

Jens Hesselbjerg Christensen,[1] Ole Bøssing Christensen, Jan-Peter Schulz

Danish Meteorological Institute

Copenhagen, Denmark

Stefan Hagemann and Michael Botzet

Max-Planck Institute for Meteorology

Hamburg, Germany

[1]Corresponding author. Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen Ø, Denmark, Phone: +45 39 15 75 00, Fax: +45 39 15 74 60, E-mail: jhc@dmi.dk

Abstract

Recently global data sets of surface elevation and distribution of major ecosystem types have been made publicly available by the U.S. Geological Survey. These data have been derived from various sources, but with a strong influence from satellite based information. From these data sets we construct a new physiographic data base of surface characteristics relevant for high resolution atmospheric models, here examplified for use in the HIRHAM4 regional climate model. The data base comprises orography and related parameters (sub-grid scale variance etc.) including roughness length, but also parameters related to the ecosystem types, such as albedo, leaf area index, and vegetation ratio. Examples are provided for a specific model configuration covering the entire European area at around 50 km horizontal resolution.

1. Introduction

Various parameters related to the characteristics of land surfaces appear in the parametrization of momentum, energy and moisture fluxes at the atmosphere-ground interface in numerical atmospheric models. Hence, a detailed knowledge of physiographic conditions plays a major role at the lower boundary of atmospheric models if realistic simulations of surface-atmosphere interactions should be made. In the regional climate model HIRHAM4 (Christensen *et al.*, 1996), developed in collaboration between the Danish Meteorological Institute, the Max-Planck-Institut für Meteorologie, Germany and the Royal Dutch Meteorological Institute, such parameters are the fractional land coverage f_l, surface mean orography $\bar{h}$, sub-grid scale orographical variance σ_h^2, and sub-grid scale directional variances $\sigma_{N-S}^2, \sigma_{NE-SW}^2, \sigma_{E-W}^2$, and σ_{SE-NW}^2. The latter are used in the gravity wave drag formulation applied in the model. Moreover, the background albedo α (the albedo of snow-free land surfaces), the surface roughness length z_0, which is constructed from a vegetation related term, z_{veg}, and a term stemming from sub-grid scale orographical variance z_{oro} (i.e. $z_0 = \sqrt{z_{oro}^2 + z_{veg}^2}$), the leaf area index LAI, the fractional vegetation cover c_v, the forest ratio (or the fractionally forrested area of a gric box) c_F (solely used to modify the albedo over snow-covered forested areas), and the maximum soil water holding capacity w_{Smax} need to be specified. In addition, a soil type is needed for determination of additional soil characteristics as required by the surface scheme (heat capacity and conductivity). At present the soil maps from FAO/Unesco (1981) at a $0.5^\circ$by 0.5°resolution is used. In HIRHAM4 these parameters have so far been treated as constant in time. Other parameters like surface temperature, soil water content, and snow depth are treated as prognostic variables in the model, and are therefore computed during the simulation. Sea surface temperature and sea ice coverage are prescribed and usually provided once a day in a typical model application. However, even for climate purposes an initial model state also for these parameters is required. Mean values of these latter parameters with a temporal resolution of one month have for practical purposes been specified in the climate data base for the HIgh Resolution Limited Area Model (HIRLAM) system (Källén, 1996).

A climate and physiography data base was developed as a part of the HIRLAM forecasting system (Machenhauer, 1988). Essentially, this data base was a derivative of an early version of the global Euro-

pean Centre for Medium-Range Weather Forecasts (ECMWF) Climate System (Brankovic and Maanen, 1985), serving a similar purpose for the ECMWF forecasting system at that time. The main physiographic input data set to this system is the US Navy global orography data set with a horizontal resolution of 10 nautical miles (18 km). For the global climate model ECHAM4 (Roeckner *et al.*, 1996) and successively for HIRHAM4, an additional data set of land-surface parameters at a resolution of 0.5°times 0.5°has been designed (Claussen *et al.*, 1994). Neither of these data sets are of a sufficient resolution for simulations at finer scale (below approx. 50 km). While the parameters calculated from the land use classification have been introduced in Hagemann *et al.* (1999), the present document mainly serves to document the parameters derived from the new elevation data set. We do, however, also describe a newly introduced seasonality of some of the ecosystem deduced variables.

2. Overview of the former HIRHAM climate file system

The former HIRHAM climate file system has been based on the modified ECMWF system adjusted to suit needs in the HIRLAM numerical weather prediction model as described in the HIRLAM documentation (Källén, 1996), with later amendments to allow for the additional fields required in the ECHAM4 physics (Roeckner *et al.*, 1996). For further detail about the underlying climatological data and procedures to establish the ECMWF climatological fields, the reader is referred to Brankovic and Maanen (1985).

The following computational steps were carried out in the old climate file system:

1) Generation of mean surface elevation and sub-grid scale standard deviation, fraction of land or sea and orography-dependent roughness on the required HIRHAM grid by summation over all input (US Navy data base at 10 arc min resolution) gridpoints within each HIRHAM grid square. Since a summation over the input data is carried out only within each HIRHAM grid square, the technique is not applicable for gridsizes below approximately 20 km.

2) Directional variances were interpolated from T213 (equivalent to approx. 60 km grid point spacing) data available at ECMWF, referring to the operational model configuration for the mid 1990's.

3) Generation of background albedo, the vegetation dependent part of the roughness length, leaf area index, fractional vegetation cover, forest ratio, FAO soil type, and maximum soil water holding capacity

on the required HIRHAM grid either by summation over grid point values or by bilinear horizontal interpolation of the data set of Claussen *et al.* (1994), which has a horizontal resolution of 0.5°by 0.5°. The resolution of this data set makes it inadequate for gridsizes finer than approximately 50 km.

4) Interpolation of monthly mean surface temperatures from an older global coarse mesh (5°by 5°) data set and adjustments of these temperatures to the model orographical grid point height. These surface temperatures are used to estimate initial soil temperatures (5 levels) by a solution of the heat conduction equation for an annual cycle.

5) Interpolation of monthly mean precipitation amounts from a global coarse mesh data set (5°by 5°). Derivation of monthly mean soil moisture fields and snow depth fields from the precipitation and surface temperature fields. This soil moisture is subsequently scaled to the maximum soil water holding capacity during the very first prepatory settings when running the HIRHAM model.

6) Interpolation of monthly mean sea surface temperature fields from a global sea surface temperature data set (Alexander and Mobley, 1971). From these data the corresponding fields of fractional sea ice cover are generated.

7) Blending of the vegetational roughness part with the orography based part by quadratic summation ($z_0 = \sqrt{z_{oro}^2 + z_{veg}^2}$).

8) Application of a weak Gaussian spatial filtering of the orography and surface roughness fields to reduce numerical noise near steep topographical gradients.

In HIRHAM4 little attention is paid to the initialisation of the seasonally varying fields that are calculated by the model itself as the prognostic equations are solved. This includes fields such as surface temperature, soil moisture and snow coverage. These fields are very difficult to ascribe realistically and it is supposed that the model will generate its own climate, which more adequately will reflect the impact of the high resolution. In Christensen (1999), however, it was shown that it is by no means a trivial task to initialise the soil components of the HIRHAM4 model satisfactorily, and a procedure on how to improve the initial model state is put forward.

3. The new physiographical data sets

a. Topography

The availablity of global data sets of elevation information was hindered for many years for political reasons. With the changes on the world political scene during the 1990's, such global information became gradually more accessible. This resulted in the production of the gtopo30 data base. The gtopo30 database, completed in late 1996, was developed over a three year period through a collaborative effort led by staff at the U.S. Geological Survey's EROS Data Center. gtopo30 is a global digital elevation model with a horizontal grid spacing of 30 arc seconds (approximately 1 kilometer). It was derived from several raster and vector sources of topographic information. Detailed information on the characteristics of gtopo30 including the data distribution format, the data sources, production methods, and accuracy estimates, can be found in the gtopo30 README file (http://edcdaac.usgs.gov/gtopo30/README.html). Based on this data set various physiographical parameters can be deduced. Mean orography, orographical roughness length and orographical variances can easily be deduced for any HIRHAM model grid by aggregation of the high resolution data to the model target grid, which is typically at a resolution between 15 and 50 km.

In constructing the mean orography $\bar{h}$ within a target gridbox a summation over each gtopo30 gridbox value within the target grid is performed, e.g.

$$\bar{h} = \frac{1}{n}\sum_{i=1}^{n} h_i \tag{1}$$

Supplementry to the elevation, the variance σ_h^2 of this field is constructed

$$\sigma_h^2 = \frac{1}{n-1}\sum_{i=1}^{n} (h_i - \bar{h})^2 \tag{2}$$

This variable is needed for the parameterization of runoff (Dümenil and Todini, 1992). Furthermore, this variable is introduced in the formulation of the sub-grid scale orographical variance z_{oro}. The formulation we have adopted follows the work proposed by Tibaldi and Geleyn (1981) with certain modifications. These have been motivated by the fact that the gtopo30 data set resolves all major obstacles and a parameterization of unresolved peaks can be omitted. Hence,

$$z_{oro} = \sqrt{\frac{n_p}{A}} \times \sigma_h^2 \tag{3}$$

where n_p is the number of significant height maxima within the target grid box and A denotes the area of the target grid box. In order to achieve realistic values (more precisely: a reasonable agreement with the old data base), a threshold has been introduced defining a maximum to be significant, only if the peak is more than 50 m heigher than the surrounding gtopo30 heights. This procedure is far from optimal, and alternatives are highly desirable; see also Sattler (1999).

The sub-grid scale directional variances $\sigma^2_{N-S}, \sigma^2_{NE-SW}, \sigma^2_{E-W}$, and σ^2_{SE-NW} are constructed as follows: Within each target grid, all the gtopo30 data points are located within $N \times N$ sub-grid boxes. Following this, a summation along perpendicular directions (in the following denominated by i and j, respspectively) is performed. This way the variances are defined as:

N-S variance:

$$\sigma^2_{N-S} = \frac{1}{2N-1} \sum_{j=1}^{N} (\bar{h_j} - \bar{h})^2; \qquad \bar{h_j} = \frac{1}{N} \sum_{i=1}^{N} h_{i,j} \tag{4}$$

NE-SW variance:

$$\sigma^2_{NE-SW} = \frac{1}{2N-1} \sum_{k=1}^{2N-1} (\bar{h_k} - \bar{h})^2; \qquad \bar{h_k} = \begin{cases} \frac{1}{k} \sum_{i=1}^{k} h_{i,k+1-i} & 1 <= \text{k} <= \text{N} \\ \frac{1}{2N-k} \sum_{i=k-N+1}^{N} h_{i,k+1-i} & \text{N} <= \text{k} <= \text{2N-1} \end{cases} \tag{5}$$

E-W variance:

$$\sigma^2_{E-W} = \frac{1}{2N-1} \sum_{i=1}^{N} (\bar{h_i} - \bar{h})^2; \qquad \bar{h_i} = \frac{1}{N} \sum_{j=1}^{N} h_{i,j} \tag{6}$$

SE-NW variance:

$$\sigma^2_{SE-NW} = \frac{1}{2N-1} \sum_{l=1}^{2N-1} (\bar{h_k} - \bar{h})^2; \qquad \bar{h_k} = \begin{cases} \frac{1}{k} \sum_{i=1}^{k} h_{i,N+i-k} & 1 <= \text{k} <= \text{N} \\ \frac{1}{2N-k} \sum_{i=k-N+1}^{N} h_{i,N+i-k} & \text{N} <= \text{k} <= \text{2N-1} \end{cases} \tag{7}$$

b. Land use classification

A 1 km global data set of major ecosystem types according to Olson (Olson, 1994a; Olson, 1994b) including glacial ice and open water has recently been made available also by the U.S. Geological Survey (USGS, 1997). It has been derived from the International Geosphere Biosphere Programme (IGBP) 1 km AVHRR data set. In Hagemann *et al.* (1999) it is documented how a global data set of land surface parameters is constructed by allocating parameters to each ecosystem class. The parameters are:

background surface albedo, surface roughness length due to vegetation, fractional vegetation cover and leaf area index for the growing and the dormancy season, forest ratio, plant-available soil water holding capacity, and volumetric wilting point. This global data set is provided for use in both the HIRHAM and the global ECHAM4/5 models. Table 2 in Hagemann *et al.* (1999) summarizes the parameter look-up values based on the Global Ecosystem Legend of (Olson, 1994a; Olson, 1994b).

Furthermore, a seasonal variation of the albedo, *LAI* and vegetation ratio has been introduced, where the monthly anomalies have been deduced from Normalized Difference Vegetation Index (NDVI) based satellite imagery (see also section 4d). This seasonality is added after the high resolution data has been processed, e.g. at the HIRHAM target grid.

4. Comparison with old data set

In this section we compare the HIRHAM physiographical data used in the older version with the new data for a domain that has been established for model development over the European continent in the EU-project "Modelling European Regional Climate, Understanding and Reducing Errors" (MERCURE).

The adopted model configuration is determined by the computational grid utilising a rotated longitude latitude coordinate system, with the coordinates of the rotated South Pole at 27 °E, 37 °S, whereby the rotated equator crosses the middle of the model domain, minimizing projection effects. The grid mesh has 110 longitudinal points and 104 latitudinal points. The coordinates in the rotated system of the SW corner are (32.65 °W, 22.62 °S) and the NE corner is at (15.31 °E, 22.70 °N).

a. Orography

In Fig. 1 the old and new surface orography is displayed in the top panels, while the difference is shown below. Major differences are identified over the Scandinavian Penisula in that the Norwegian mountains are higher in the new data set. In general there are many differences in steep topographical terrain. The new data set tends to sharpen the mountain ranges. Note in particular major changes over Scandinavia and Iceland as well as along the Adriatic coast line. Moreover, the elevation has generally been slightly increased over low land areas for major parts of Europe.

b. Orographical variances

Figure 2 shows the orographical variance using the old data set and the new one, respectively. Here, the high resolution of the new data set has a very remarkable impact in all mountaineous regions, with resulting major increases of the values nearly everywhere. As a consequence, the simulated surface runoff is enhanced in these regions and less water is stored in the soils.

Figures 3 to 6 compare the directional variances from the old and new source, respectively. Again substantial differences are identified. In fact, the old interpolated data were not really representing the necessary resolution, which is also visible in many cases.

c. Rougness length

Figure 7 compares the old and new estimated roughness lengths. Noticable differences occur in the mountains as the procedure to account for topographical peaks in the data set has been substantially modified. But also the introduction of the thresholding parameter in counting peaks has some consequences. This basically reduces orographically induced roughness over the low land areas to zero, while the vegetational part at the same time has been dramatically altered (see (Hagemann *et al.*, 1999)). Hence, we are not convinced that any of the roughness length maps shown are adequate and particularly representative in mountaineous regions.

d. Other parameters

As mentioned before, Hagemann *et al.* (1999) have constructed a global data set of land surface parameters which is based on a 1 km global distribution of major ecosystem types (according to Olson (1994a)) including glacial ice and open water that was recently made available by the U.S. Geological Survey (USGS, 1997). This global data set is available for use in regional and global climate modelling, and it is implemented in both the HIRHAM and the new version 5 of the ECHAM global climate modelling system. The parameters included are: background surface albedo, surface roughness length due to vegetation, fractional vegetation cover and *LAI* for the growing and dormancy season, forest ratio, plant-available and total soil water holding capacity. These parameters represent annual mean values of land surface characteristics.

Here, we simply compare the mean fields for completeness. Figure 8 compares albedo fields, Figure 9 forested fraction, Figure 10 vegetation, and Figure 11 the leaf area index. Finally, Figure 12 compares the new and old maximum soil water holding capacities.

In addition, the seasonal variation of vegetation characteristics is provided. Monthly mean fields of vegetation ratio, leaf area index and background albedo were developed and implemented. The seasonal cycles of vegetation ratio and *LAI* are determined from their minimum and maximum values given by Hagemann *et al.* (1999), together with monthly values of fraction of absorbed photosynthetically active radiation *FPAR* (Knorr, 1998), and a monthly 2m temperature climatology TLW (Legates and Wilmott, 1990). The data sets of *FPAR* and TLW were used to define a global field of a monthly growth factor f_i which determines the growth characteristics of the vegetation at a resolution of 0.5°by 0.5°. This growth factor represents the local climate and it does not vary largely between grid boxes in a certain climate region. Thus, it can be aggregated to coarser resolutions and may also be applied to data with higher resolutions than 0.5°without a significant loss of information. Since *FPAR* is a direct measure of the amount of vegetation on the land surface, it is used for the definition of the growth factor f_i in all grid boxes where a *FPAR* value is available in all 12 months of the year. This comprises merely the low and mid-latitudes, while in high latitudes, no *FPAR* values are available throughout the year due to snow coverage and the course of the satellite orbit. For a certain grid box, f_i is defined by

$$f_i = 1 - \left(\frac{FPAR_{max} - FPAR_i}{FPAR_{max} - FPAR_{min}} \right)^2 \tag{8}$$

$FPAR_i$ is the *FPAR* value of month *i*, $FPAR_{min}$ and $FPAR_{max}$ are the minimum and maximum monthly *FPAR* values for the grid box, respectively. In high latitudes, the growth of vegetation is mainly limited by temperature. Here, TLW is used to define f_i in the following way

$$f_i = 1 - \left(\frac{T_{max} - T_i}{T_{max} - T_{min}} \right)^2 \tag{9}$$

T_i is the climatological 2m temperature of month *i*. In high latitudes, it is assumed that the minimum vegetation is present for temperatures below T_{min} = 278 K and the maximum vegetation is present at the maximum monthly 2m temperature T_{max} or for temperatures above T_{max} = 298 K, whichever is lower. Using the growth factor f_i based on the Eqs. (8) and (9), the monthly LAI_i (analogous for the vegetation

ratio) can be computed as

$$LAI_i = LAI_{min} + f_i(LAI_{max} - LAI_{min}) \tag{10}$$

The background albedo α_i of month *i* was derived from the LAI_i of month *i* based on Zeng *et al.* (1999):

$$\alpha_i = \alpha_0 - c(1 - \exp(-0.5LAI_i)) \tag{11}$$

c is set to a default value of 0.15 and α_0 is computed from Eq. (11) inserting the annual averages of albedo and *LAI* obtained by Hagemann *et al.* (1999). Both *c* and α_0 are corrected to achieve minimum albedo values (corresponding to the maximum *LAI*) of 0.09 (Zeng, personal communication, 2000) where the annual mean albedo is above 0.09, thereby retaining the same annual mean albedo.

The monthly variations of vegetation ratio, *LAI* and background albedo are shown in Figures 13 to 15, respectively. The seasonal cycle of deciduous forest with low values of vegetation ratio and *LAI* in winter and high values in summer (and vice versa for surface albedo) is well captured by the new data sets, as clearly seen in central Europe. On the other hand, needleleave trees in parts of Scandinavia like Sweden and Finland show only low variations of these vegetation related parameters, which is also reasonable. It is the hope that with these new parameter sets, capturing important features of different vegetation types, we get a more realistic representation of vegetation and land surface in general in climate model simulations.

Acknowledgements

This work was supported by the EC Environment and Climate Research Programme (contracts: ENV4-CT97-0485 (MERCURE) and ENV4-CT97-0522 (TUNDRA), Climate and Natural Hazards).

Alexander, R. C., and R. L. Mobley, 1971: Monthly average sea-surface temperatures in ice-pack limits for a 1 degree global grid. Report R-1310-ARPA, RAND.

Brankovic, C., and J. V. Maanen, 1985: The ECMWF climate system. ECMWF Report 109, ECMWF.

Christensen, J. H., O. B. Christensen, P. Lopez, E. van Meijgaard, and M. Botzet, 1996: The HIRHAM4 regional atmospheric climate model. Scientific Report 96-4, Danish Meteorological Institute.

Christensen, O. B., 1999: Relaxation of soil variables in a regional climate model. *Tellus*, **51A**, 674–685.

Claussen, M., U. Lohmann, E. Roeckner, and U. Schulzweida, 1994: A global data set of land-surface parameters. MPI Report 135, Max-Planck Institut für Meteorologie.

Dümenil, L., and E. Todini, 1992: A rainfall-runoff scheme for use in the Hamburg climate model. In J.P. O'Kane (Ed.), *Advances in Theoretical Hydrology*, Volume 1 of *EGS Series on Hydrological Sciences*, pp. 129–157. Elsevier.

FAO/Unesco, 1981: Soil map of the world. Vols. 1-10, Unesco.

Hagemann, S., M. Botzet, L. Dümenil, and B. Machenhauer, 1999: Derivation of global GCM boundary conditions from 1 km land use satellite data. MPI Report 289, Max-Planck Institut für Meteorologie.

Källén, E. (Ed.), 1996: *HIRLAM Documentation Manual, System 2.5*. The Swedish Meteorological and Hydrological Institute. Available from SMHI, S-60176, Norrköping, Sweden.

Knorr, W., 1998. Constraining a global mechanistic vegetation model with satellite data. In: 'IEEE International Geoscience and Remote Sensing Symposium' Proceedings, Seattle, WA, 6-10 July 1998.

Legates, D., and C. Wilmott, 1990: Mean seasonal and spacial variability in global surface air temperature. *Theor. Appl. Climatol.*, **41**, 11–21.

Machenhauer, B., 1988: The HIRLAM final report. , HIRLAM Technical Report No. 5, 116pp.

Olson, J. S., 1994a: Global ecosystem framework-definitions. , USGS EROS Data Center International Report.

Olson, J. S., 1994b: Global ecosystem framework-translation strategy. , USGS EROS Data Center International Report.

Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida, 1996: The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate. Report 218, Max-Planck-Institut für Meteorologie.

Sattler, K., 1999: New high resolution physiographic data and climate generation for the HIRLAM forecasting system. DMI Technical Report 99-11, Danish Meteorological Institute.

Tibaldi, S., and J. Geleyn, 1981: The production of a new orography, land/sea mask and associated climatologicalsurface fields for operational purposes. 40, ECMWF Research Department Technical Memorandum.

USGS, 1997: Global land cover characteristics database vs. 1. , U.S. Geological Survey, http://edcdaac.usgs.gov/glcc/glcc.html.

Zeng, N., J. Neelin, W.-M. Lau, and C. Tucker, 1999: Enhancement of interdecadal climate variability in the sahel by vegetation interaction. *Science*, **286**, 1537–1540.

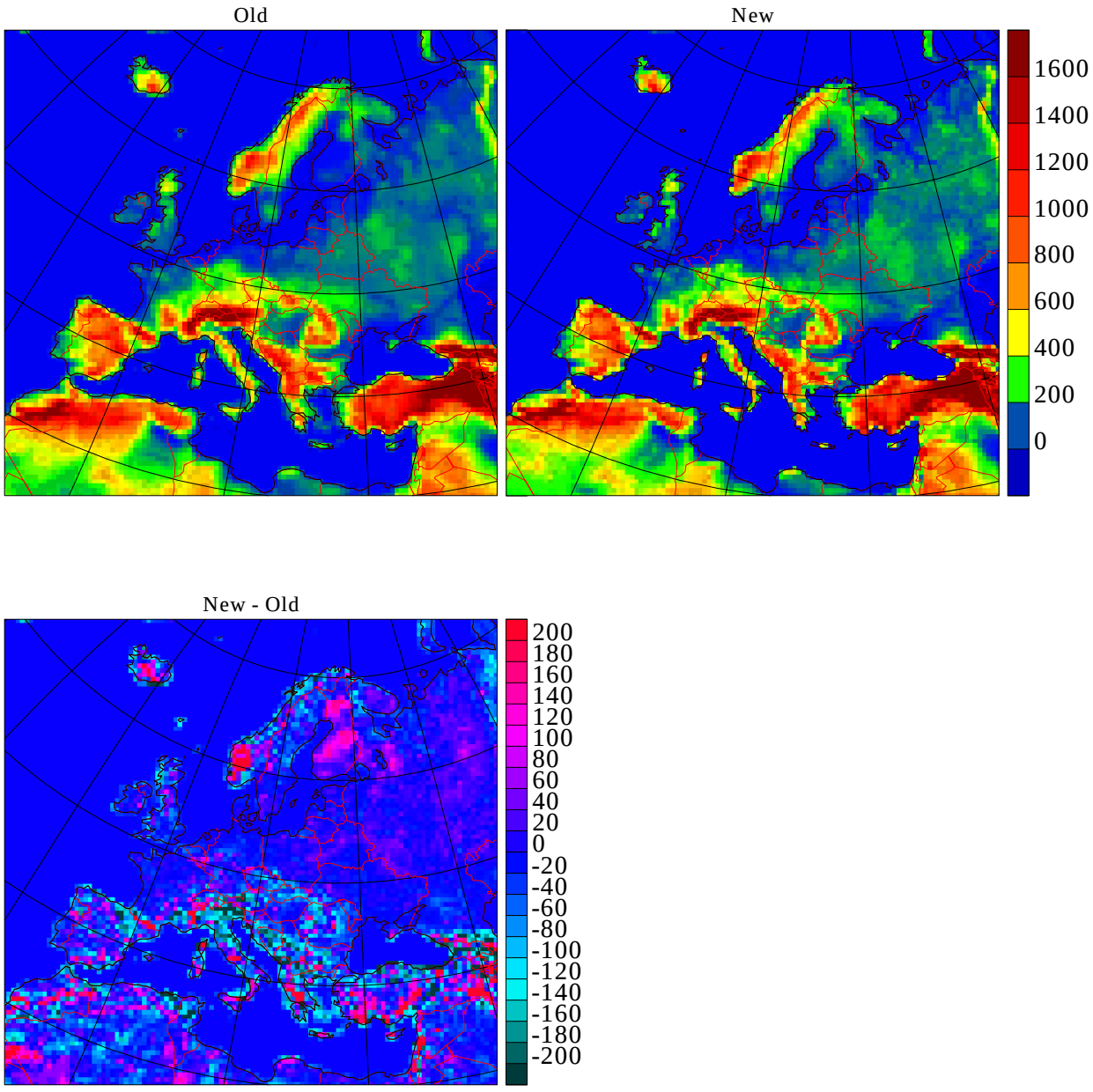


Figure 1: Old and new mean orography and their difference. Units in meters.

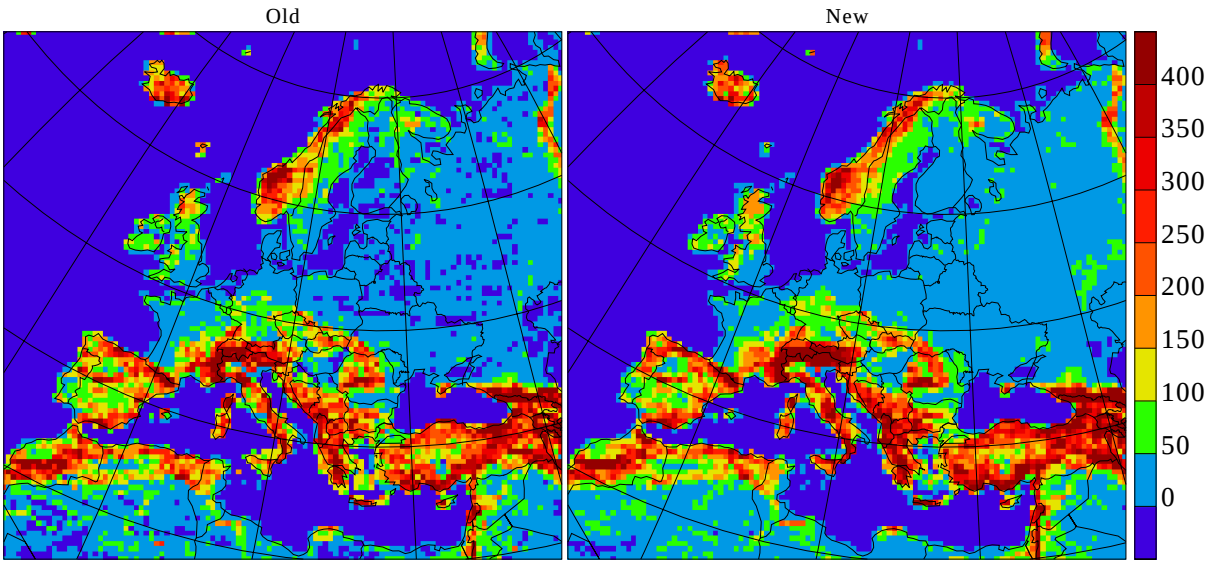


Figure 2: Square root of orographical variance. Units in meters.

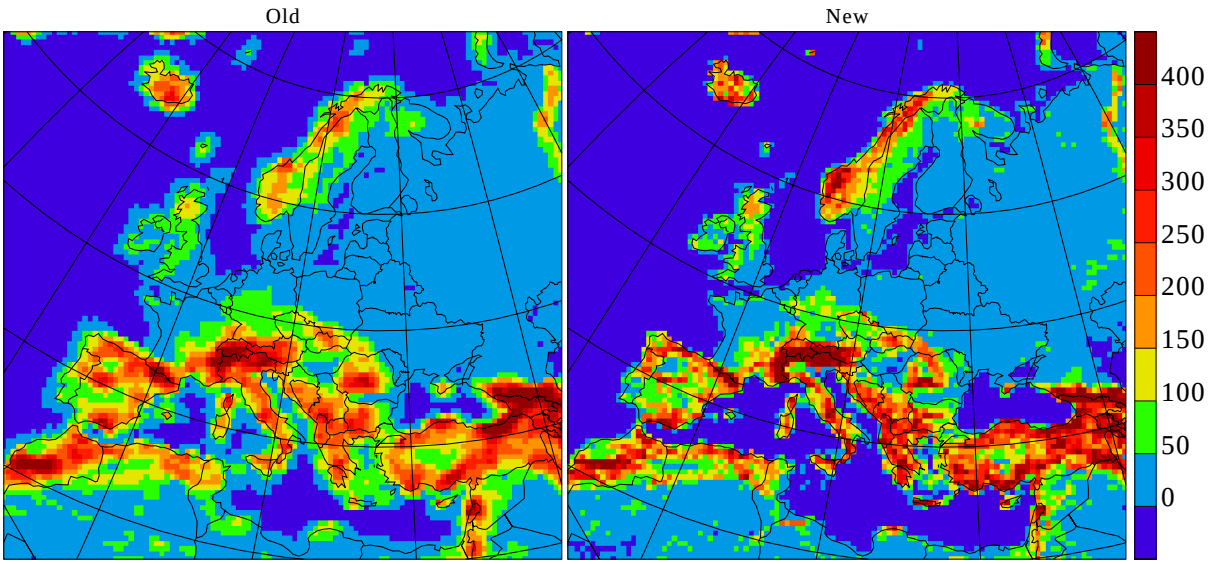


Figure 3: Square root of N-S directional variance. Units in meters.

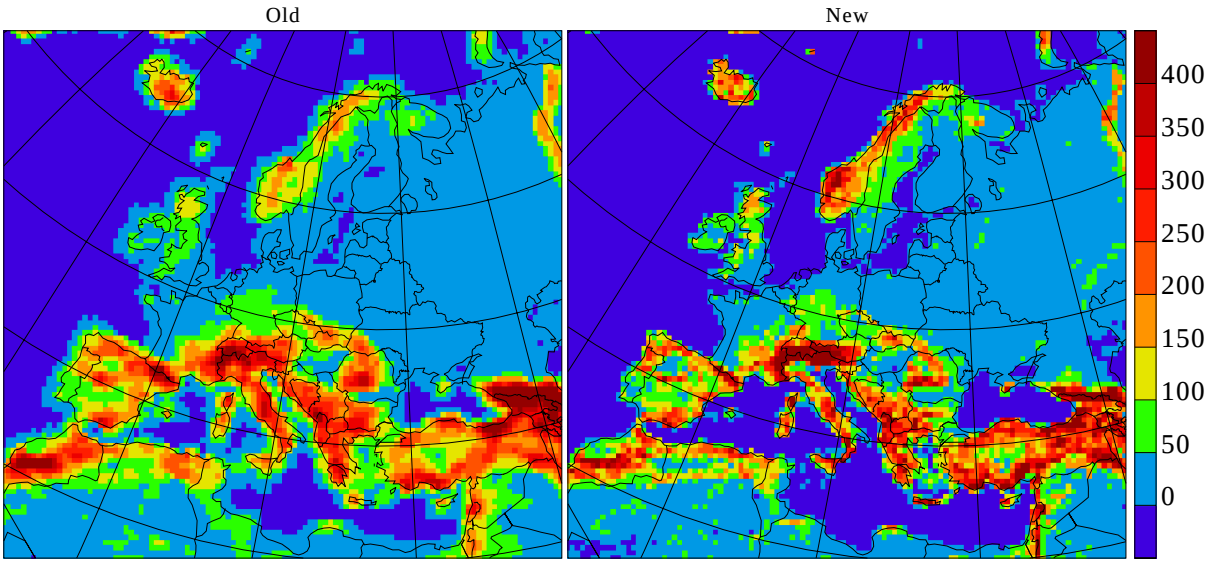


Figure 4: Square root of SE-NW directional variance. Units in meters.

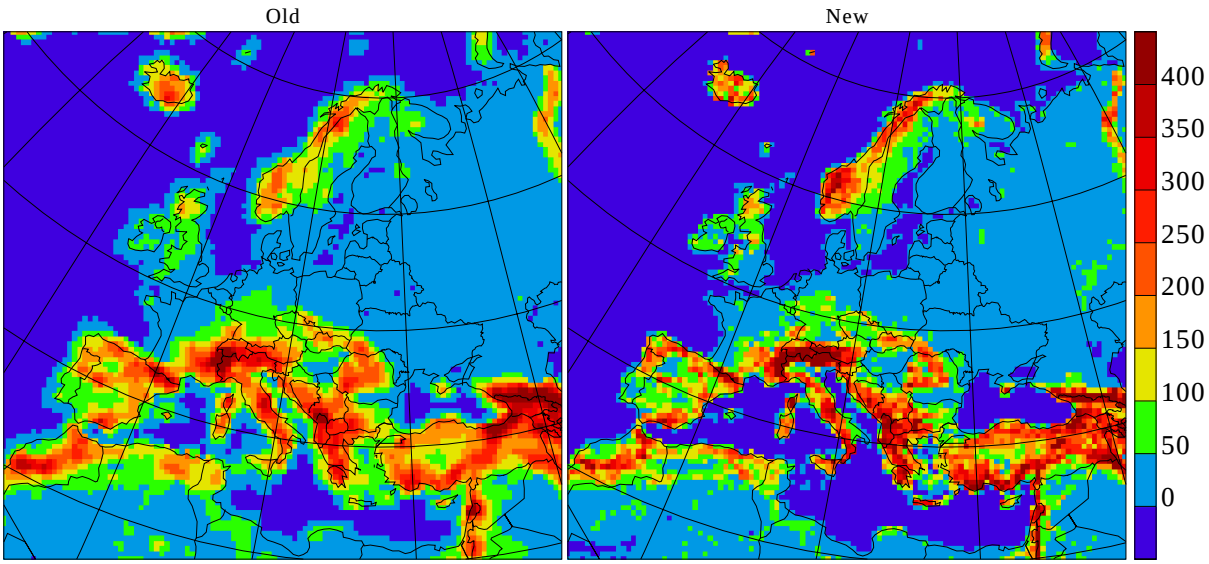


Figure 5: Square root of E-W directional variance. Units in meters.

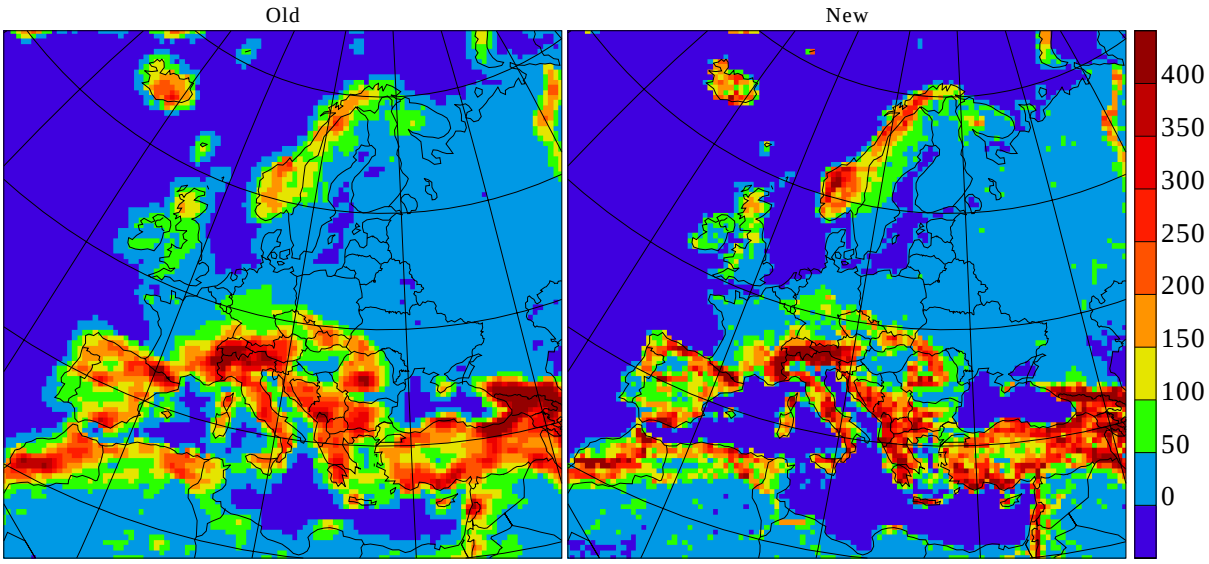


Figure 6: Square root of SW-NE directional variance. Units in meters.

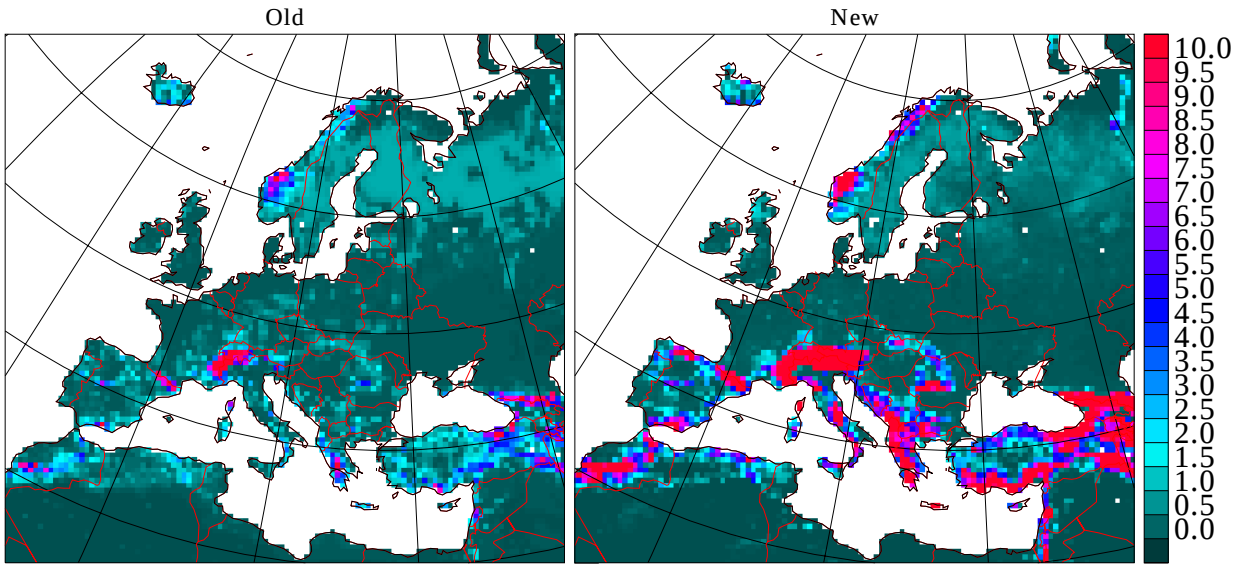


Figure 7: Roughness length. Units in meters.

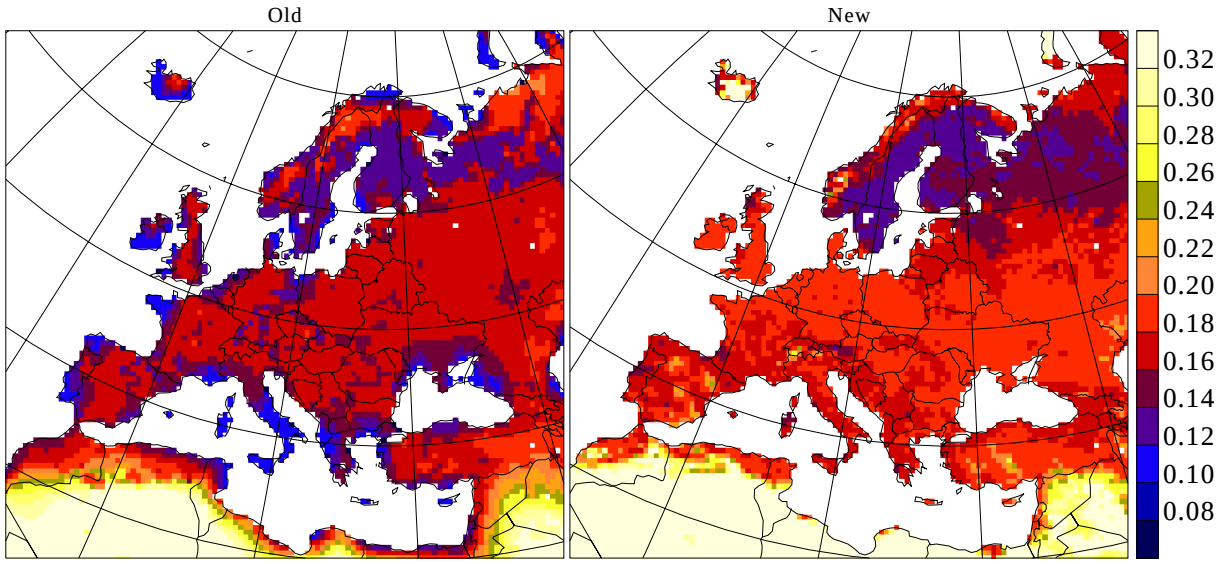


Figure 8: Annual mean background albedo.

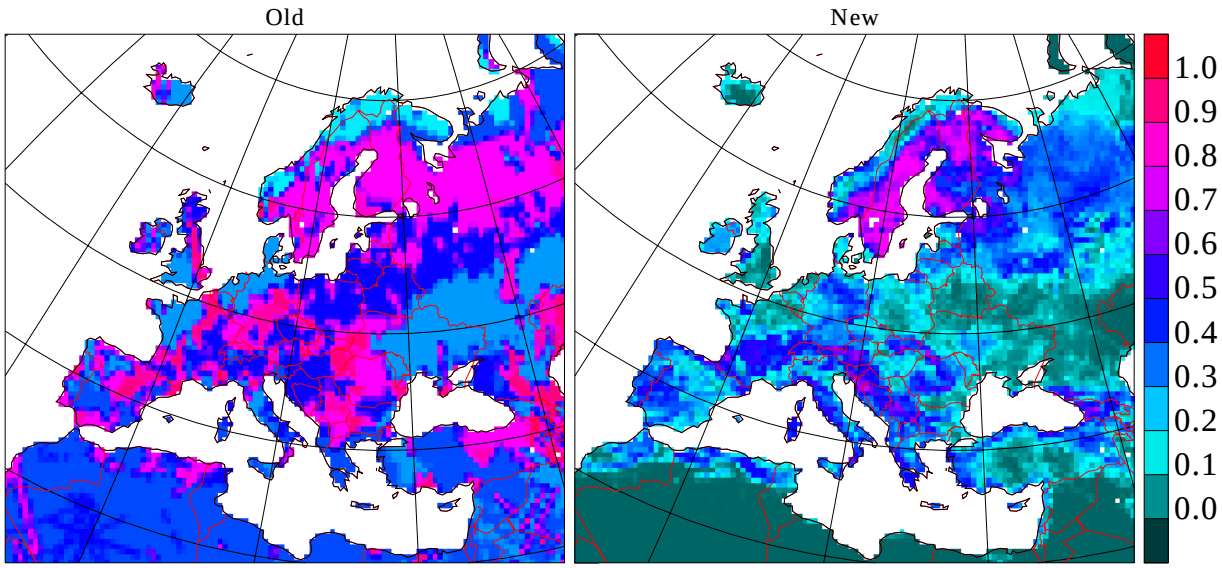


Figure 9: Forest fraction.

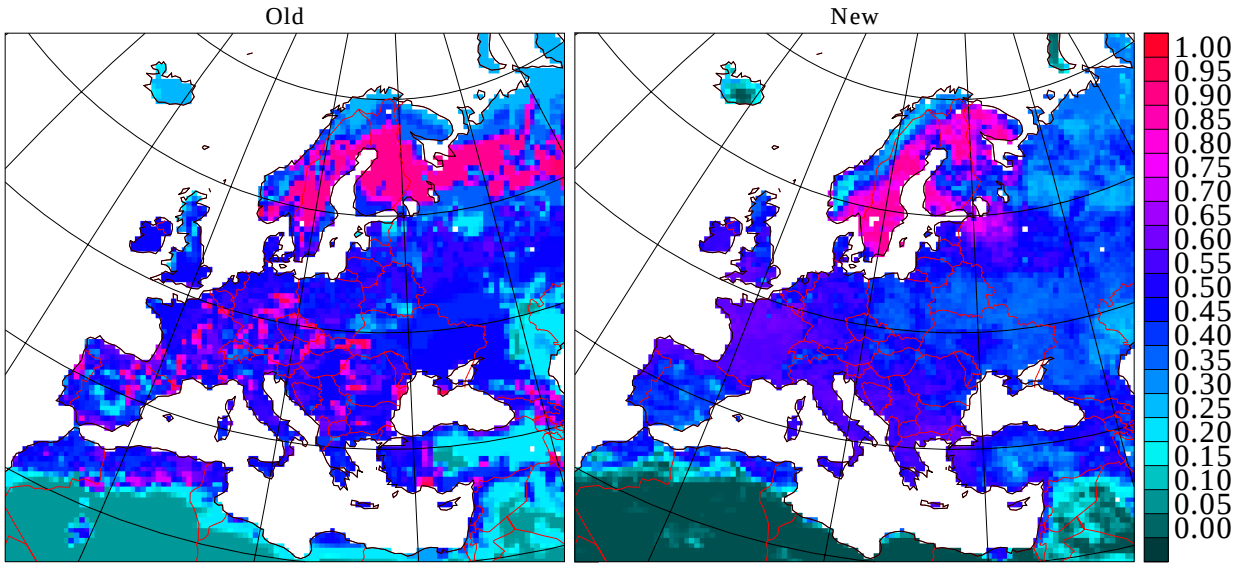


Figure 10: Annual mean vegetation ratio.

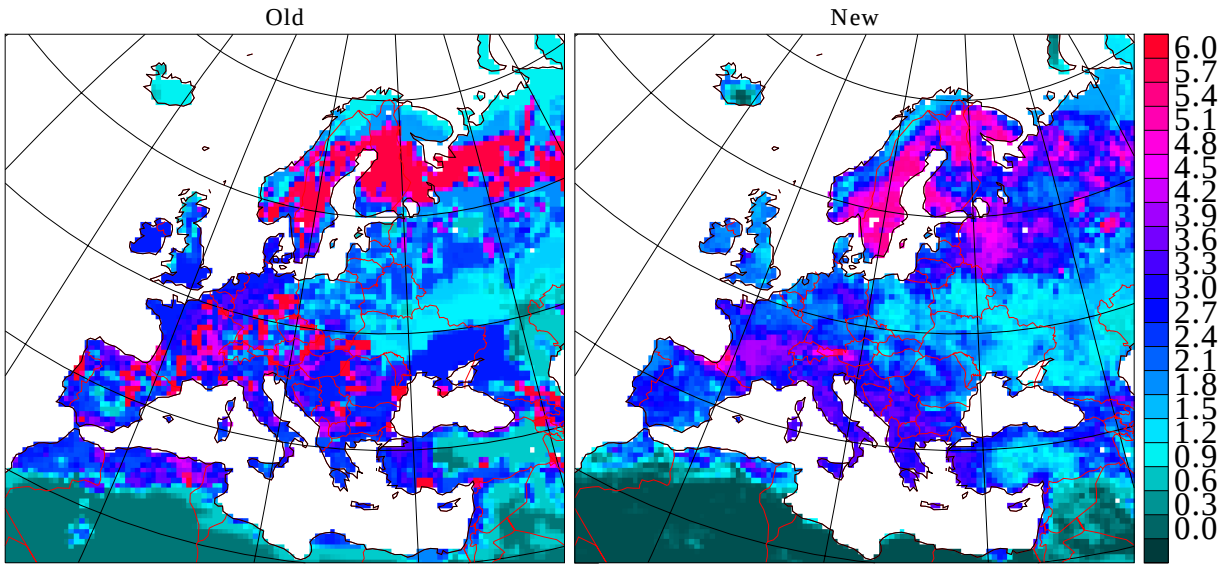


Figure 11: Annual mean leaf area index.

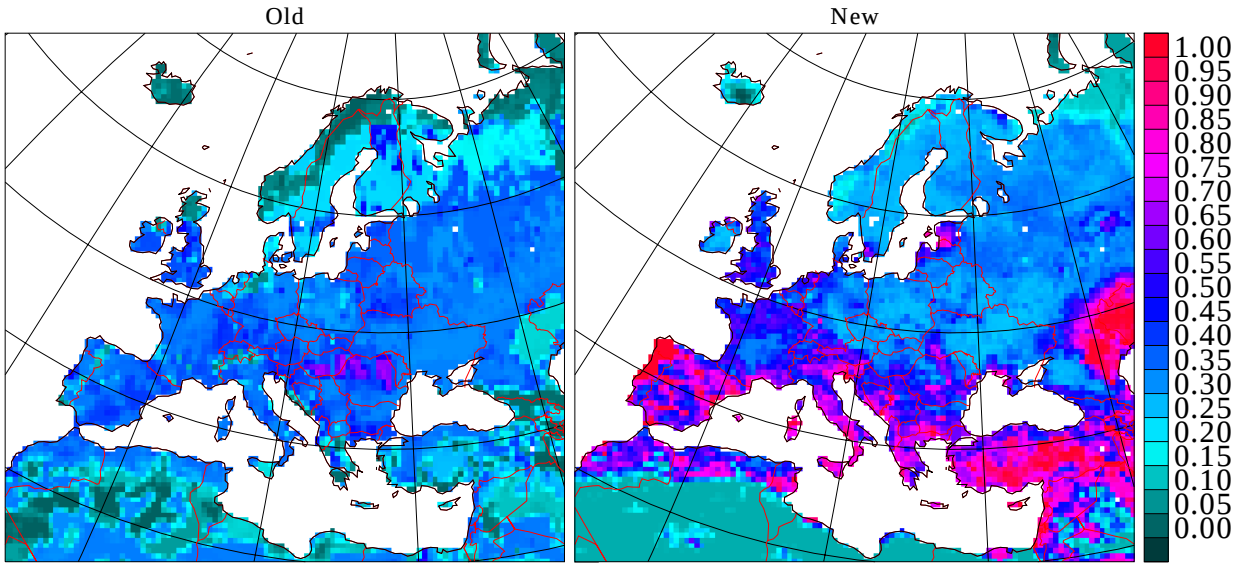


Figure 12: Maximum soil water holding capacity. Units in meters.

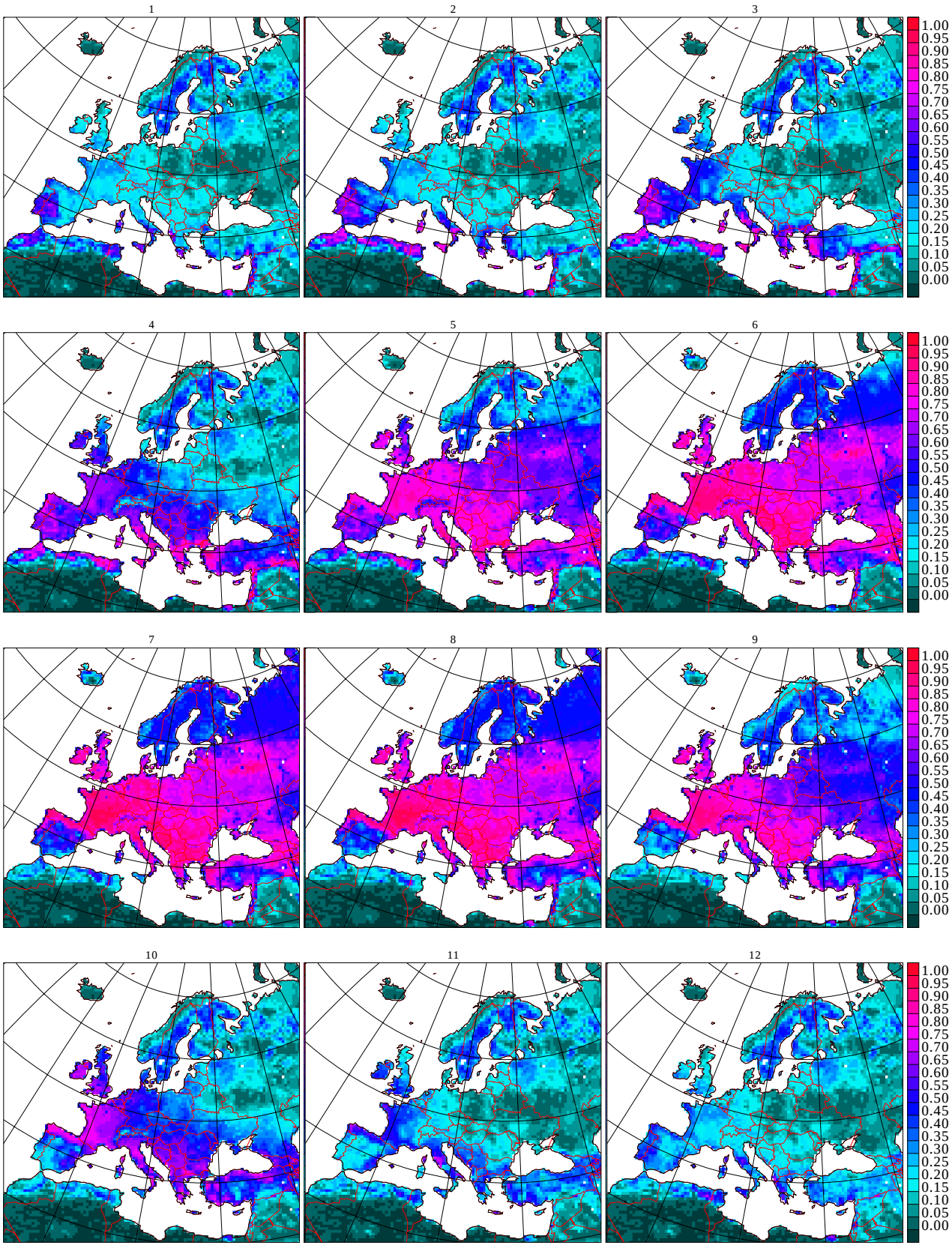


Figure 13: Seasonal variation in fractional vegetation cover.

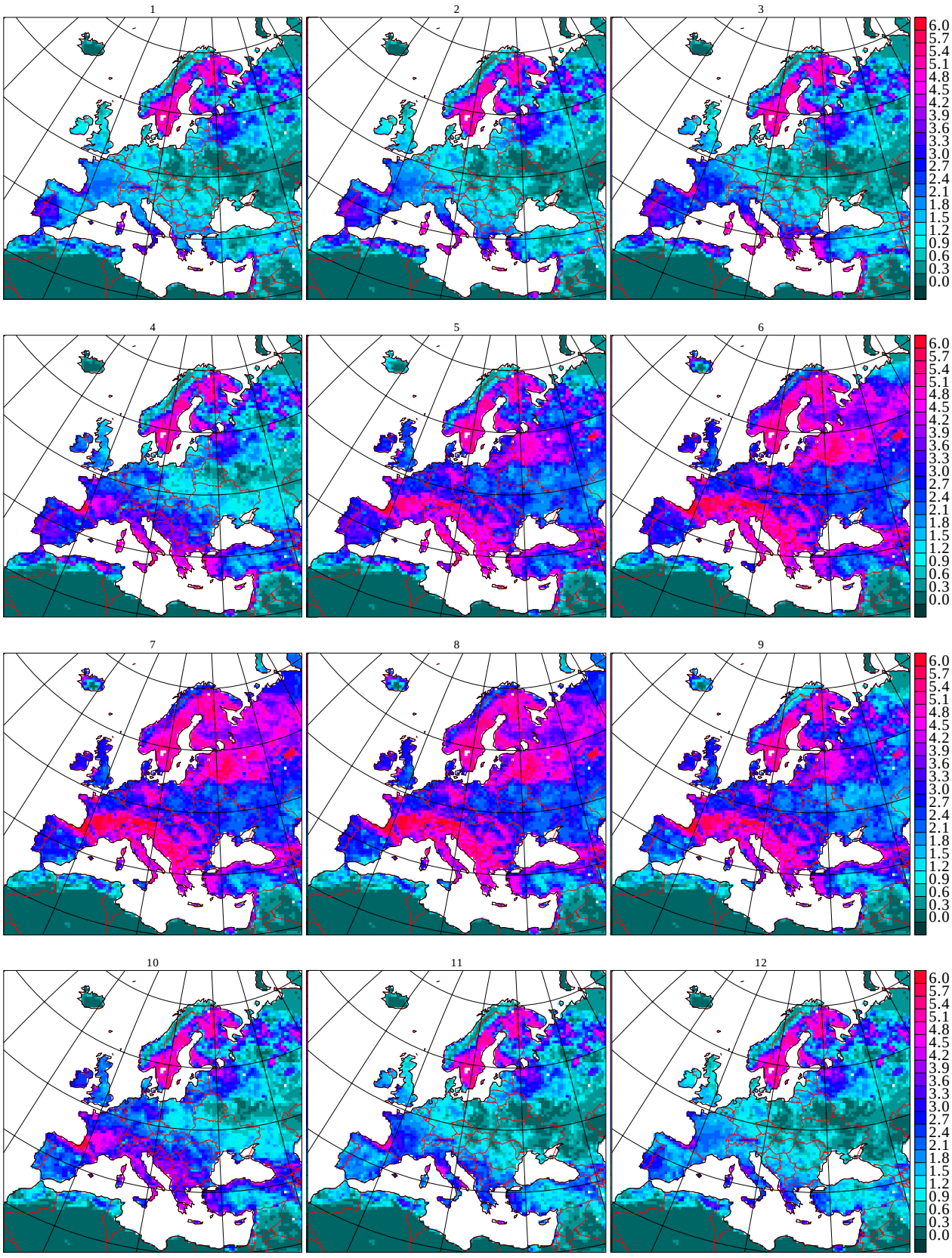


Figure 14: Seasonal variation in leaf area index.

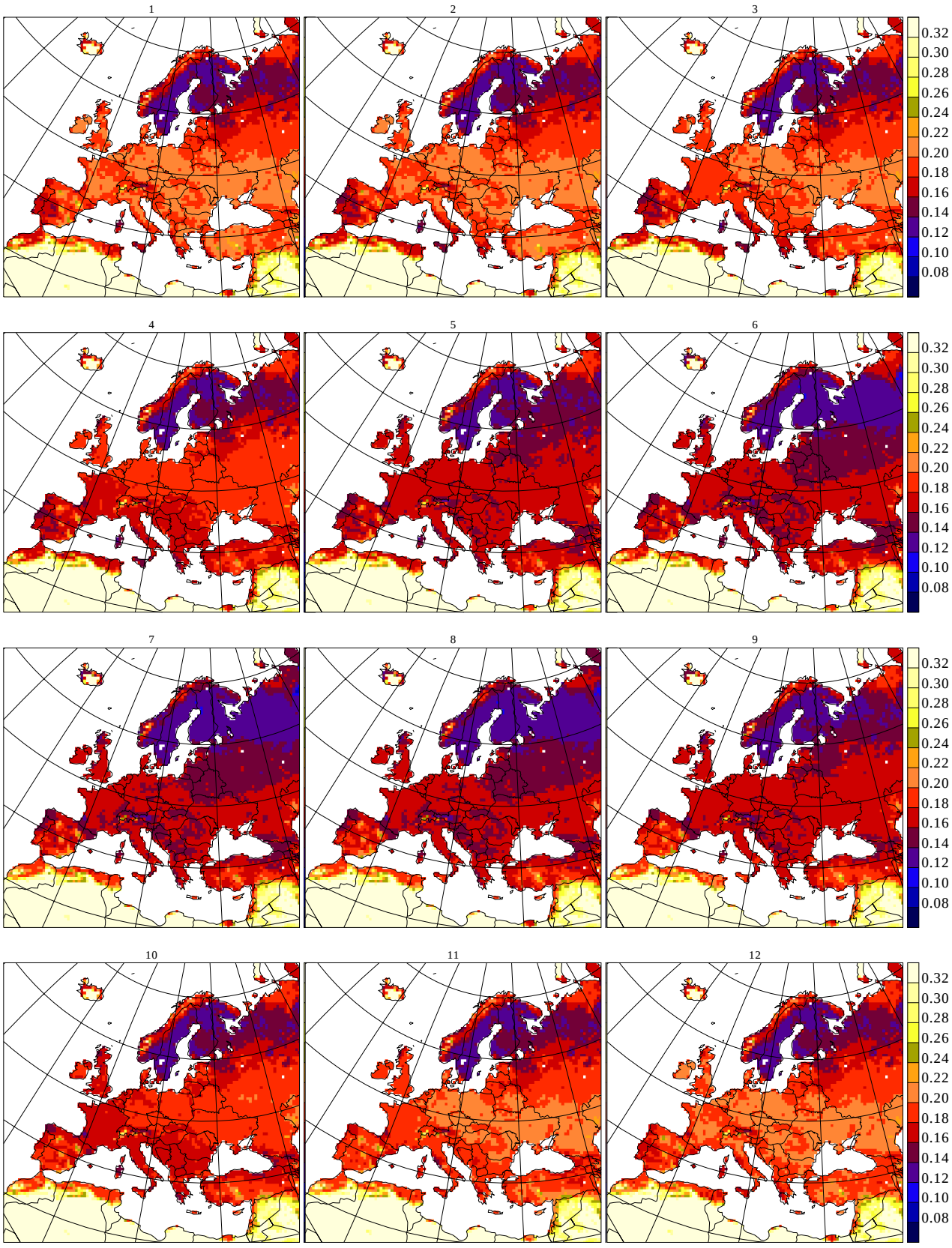


Figure 15: Seasonal variation in background albedo.