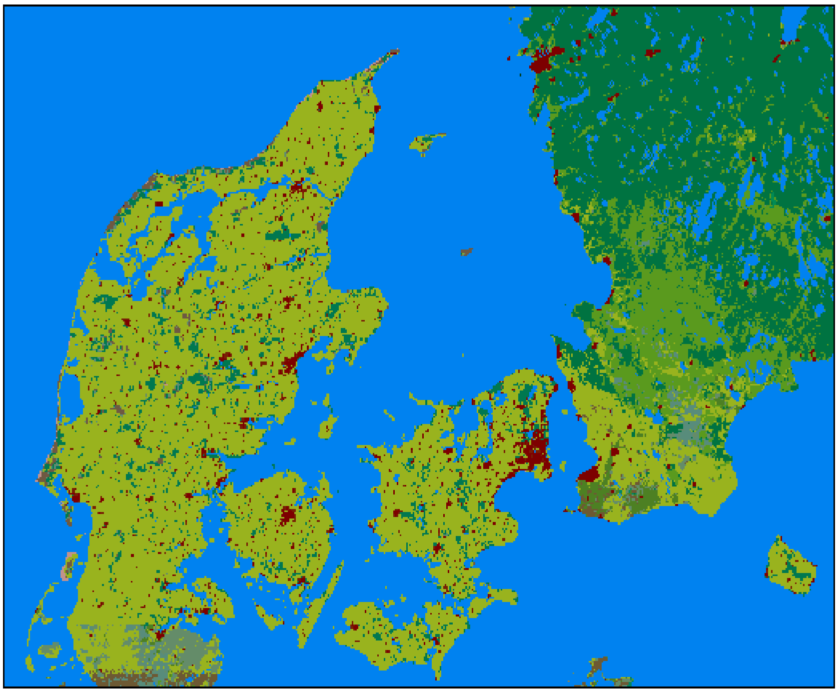
DANISH METEOROLOGICAL INSTITUTE

TECHNICAL REPORT

99-11

New high resolution physiographic data and climate generation for the HIRLAM forecasting system

Kai Sattler



DMI

COPENHAGEN 1999

ISSN-Nr. 0906-897X (printet version)
1399-1388 (online version)

1 Introduction

Physiographic data plays a major role at the lower boundary of atmospheric models like the HIRLAM system. It includes the parameters orography, land use, land–water distribution soil types and others and is represented in the model in different soil and surface properties like geopotential, roughness length, albedo etc. These properties have a direct influence upon the meteorological parameters near the ground. Their spatial distribution is often reflected in the horizontal structure of forecast fields like 2 m temperature and humidity or 10 m wind.

To improve the representation of the surface properties new high resolution physiographic data bases were introduced in the climate generation for HIRLAM at DMI. This was connected with the introduction of new software for the climate generation. It makes use of the *Hierarchical Data Format* HDF (NCSA 1998), which makes it possible to unify the format for the input files used by the climate generation and which helps to keep file sizes small. The output of the climate generation is performed in GRIB code for utilization in the operational HIRLAM at DMI (Sass 1998).

The following motivations led to using the new high resolution data bases:

- The new physiographic data is given in a horizontal resolution of approx. 1 km and is therefore much more detailed than the old data base. This is especially useful for the HIRLAM configurations with high grid resolution.
- The introduction of the ISBA (Interaction Soil–Biosphere–Atmosphere) parameterization scheme into HIRLAM (Bringfelt 1996) demands new parameters in a high resolution which cannot be retrieved from the old physiographic data.
- A remarkable improvement especially in the orography field and the roughness length is accomplished for all model domains.
- The new data bases cover the whole globe. Although not all of this data is needed and also not actually used, there is one data base for almost all model configurations.

The new data bases replace a part of the former data bases of the HIRLAM climate system (Bringfelt et al. 1995). Those parts of the former data which are still used in the new system are supplied to the climate generation in HDF-format, so that the climate generation only has to deal with one file format.

The main features of HDF are:

- many datasets can be stored in one file
- freely definable attributes can be included at different levels in order to describe the file and the datasets
- data compression capability is implemented
- there are several visualization tools to check the file contents

In order to get the data bases into HDF-format and to avoid that the complete data bases have to be addressed during the climate generation, a physiographic preprocessing is introduced. It is performed before the climate generation. Its purpose is to establish input data sets for the climate generation.

2 New Data sources

The climate generation in HIRLAM makes use of different data sources, for a part of which new high resolution data is introduced. This new data consists of two separate parts. The first part includes the *Global 30 Arc Second Elevation* data (GTOPO30) given in a regular geographical longitude-latitude grid (USGS 1998). It contains elevation heights with an accuracy of 1 meter with a spatial resolution of approx. 1 km. The second part of the new data includes the *Global Land Cover Characteristics* data (GLCC) in an equal area projection grid (USGS 1997) and a horizontal resolution of 1 km. Different data sets are available which represent different continents (e. g. Eurasia), but a global representation is also included. It is given in the *Interrupted Goode Homolosine projection* (Steinwand 1994) and was chosen as data source for the land use information. Several different land classification schemes are included in GLCC to suit the needs of different applications:

- *Global Ecosystems Legend*
- *IGBP Land Cover Legend*
- *USGS Land Use/Land Cover System Legend (Modified Level 2)*
- *Simple Biosphere Model Legend*
- *Simple Biosphere 2 Model Legend*
- *Biosphere Atmosphere Transfer Scheme Legend (BATS)*

From these schemes, the *Global Ecosystems Legend* (Olson 1994a, Olson 1994b, USGS 1997) was utilized for reasons described later on (section 3). It includes 94 land use classes (Table 1).

As the size of the original data files is enormous and not all the data from the whole globe is actually needed, the relevant part, i. e. the northern hemisphere, was cut out in a preliminary step. From this data, input files for the climate generation were created in HDF-format in a physiographic preprocessing step (section 3), which supplies the relevant data in an even more compact form than the original data bases do. The resulting files from the preprocessing include all parameters which are utilized in the climate generation.

It turned out, however, that the land use information from GLCC is not satisfactorily precise over some regions of Denmark. This deficiency occurs in all land classification schemes and especially does not suit the needs of the HIRLAM-D model. For this reason, land use data from the “Kort og Matrikelstyrelsen” (KMS) in Denmark was included to improve the land cover representation. It is originally given in a regular latitude-longitude grid with a resolution of approx. 1 km and was transformed into the *Interrupted Goode Homolosine projection*. The KMS-data consists of area fractions for bare soil, agriculture, forest, lakes, open land and urban areas. This information was transformed into land use classes of the *Global Ecosystems Legend* and included into the data base to be processed by the physiographic preprocessing. The KMS-data reveals much more details than the data from GLCC, despite of using less land use classes over Denmark (Figures 1 and 2).

3 Physiographic preprocessing

In order to avoid extreme computer resource demands during the climate generation due to the enormous size of the data bases, the physiographic data is preprocessed. This makes it possible to select only those informations from the data bases which are important in the climate generation.

The physiographic preprocessing deals with both elevation and land use data. The elevation is calculated as an arithmetic average on the demanded output grid. This may be a regular or a rotated latitude-longitude grid.

The land use data is transformed on to the same demanded grid as the elevation data. The land use classification scheme is at the same time transformed to land use classes relevant for HIRLAM and suitable for the ISBA scheme. The following land use classes are built in the physiographic preprocessing from the original *Global Ecosystems Legend*:

- cropland
- grassland short
- evergreen needle
- deciduous needle
- deciduous broadleaf
- evergreen broadleaf
- grassland tall
- desert
- tundra/wetland
- cropland irrigated
- semi-desert
- ice
- bogs
- water
- ocean
- evergreen shrubs
- deciduous shrubs
- forest mixed
- forest/fields
- urban
- field/woods

Table 1 shows the transformation between the two classification schemes. Except of the classes forest/fields and field/woods this is the classification used in the ISBA scheme (Bringfelt et al. 1995), which is very similar to the *Biosphere Atmosphere Transfer Scheme Legend* (BATS, Dickinson et al. 1986). BATS is also represented in the GLCC data base, but it is not used because it lacks a classification for urban areas and because it contains the land use class interrupted forest. This class is a very crude representation of the mixture of open areas and forest. The *Global Ecosystems Legend*, however, includes land use classes that describe different mixtures of forest and open land. They are reduced to the two categories forest/fields and field/woods within the transformation. This is still better than having only one land use class, because the two classes distinguish between a lower and a higher percentage of forest in open land.

It would of course be best to use a classification which does not contain any class representing a mixture of forest areas and open land, but GLCC unfortunately does not provide such a classification scheme, that is suitable for the purposes of HIRLAM at the same time. Nevertheless, the two retrieved classifications forest/fields and field/woods can be treated in the climate generation in a suitable way, i. e. that they represent a certain fraction of forest and a certain fraction of open land. The fraction of forest is higher in forest/fields than in field/woods (see section 4).

The land use classes listed above are stored in form of their fraction of area coverage within each field point of the output grid. So the resulting HDF file contains fields for all occurring land use classes. A land-sea mask and a land-water mask are additionally created and put into the HDF file. The latter mask resolves lakes and rivers additionally to sea water.

Supplementary to the elevation, the land use fractions and the mask fields a field containing an unscaled isotropic roughness length due to orography (z0_oro_unsc) is created. It is used in the climate generation for the determination of the sub grid orographic roughness length. A formulation based on a work by Tibaldi and Geleyn (1981) was modified an yielded realistic structures for z0_oro_unsc:

$$z_{0,oro,unsc} = \frac{1}{2}\sqrt{\frac{n_p + 0.001}{A}} var_h, \tag{1}$$

where n_p is the number of relative height maxima within the actual grid square of the demanded output grid. A denotes the area this grid square covers. The summand 0.001 is introduced to ensure a smooth field where the number of relative height maxima goes to zero. $z_{0,oro,unsc}$ depends on the resolution of the demanded output grid, because it is determined from the variance of elevation within the area of an output grid square. The variance is calculated as follows:

$$var_h = \frac{1}{n}\sum_{i=1}^{n} h_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} h_i\right). \tag{2}$$

h_i denotes height and n the number of pixels of the high resolution data in the actual grid square. The variance field is also stored in the HDF file. The data set keeping it is given a name, which contains information about the size of the grid square used for its calculation. This nomenclature identifies all variance fields unambiguously.

In order to determine a field of z0_oro_unsc that is valid for the final grid of the climate generation, it is necessary to know the geometry of the final grid already during the physiographic preprocessing. This is actually a drawback, but supplying one uniform field of sub grid orographic roughness for use in different grid resolutions is not consistent with the individual model grid.

4 The climate generation

The climate generation is the major process for supplying the physiographic and climatological parameters to HIRLAM. The introduction of the new physiographic data was also connected with technical changes in the climate generation. They refer to new routines for the complete climate generation and the exclusive utilization of the universal HDF file format. For those parameters, which are not affected by the new physiographic data, the actual output has not changed significantly.

At first, a short technical description of the routines involved in the climate generation is given. Then the input data as they are supported from the physiographic preprocessing will be shortly outlined. A description of the parameter tables defining properties for determination follows together with the steps of the climate generation. At last, the output variables will be presented and some important comments will be given.

4.1 Routines

The basic program structure of the new climate generation was taken over from KNMI (The 1998). It was modified to suit the needs at DMI. The new codes replace all of the old codes. All of the routines make use of the HDF file format. A transformation of the data into GRIB-code is performed in the last step of the climate generation. The following list outlines the programs and their purpose:

- `albedo`: Derives the monthly albedo from the albedo over land and the albedo over sea taking sea ice distribution into account.
- `clpp`: Climate Program Processor. Steering program for the climate generation.
- `ctopo`: Utility for retrieving data sets from the HDF files. The retrieval may be combined with different mathematical operations like averaging, weighted averaging, scaling and many more (The 1998). It is also possible to apply a mask field to the data and to create new data sets from several existing data sets. The retrieval of existing data sets and the creation of new data sets may be controlled through parameter tables (see below).
- `hdf2asim`: Converter for data sets from an HDF file into GRIB code.
- `sdsep`: Edit processor for HDF file data sets.
- `sharpen`: Contrast filter.
- `smooth_gauss`: Gaussian filter.
- `smooth_sha`: Shapiro filter.
- `tsoil`: Routine for calculating surface and soil temperatures, snow depth and geopotential.

`z0merge`: Combines the monthly surface roughness length for vegetation with that due to orography.

From these routines, `sdsep`, `sharpen` and `smooth_sha` are added at DMI. `clpp` was also written at DMI and is a substitute for the script `Preps`, which is supplied with the original code. `albedo` and `z0merge` are a modification of the original programs `calc_alb` and `z0combine`.

The programs are called from the *Climate Program Processor* `clpp`. It specifies the different tasks of the climate generation process and performs them by making use of a variety of parameter tables. They define meteorological properties set for the land use categories of the *Global Ecosystems Legend* (section 3) and suit the needs of the ISBA scheme. Table 2 shows how the tasks of the former climate generation routines are distributed in the new code.

4.2 Input Data

As already mentioned in the former sections, the new climate generation only uses input data given in HDF-format.

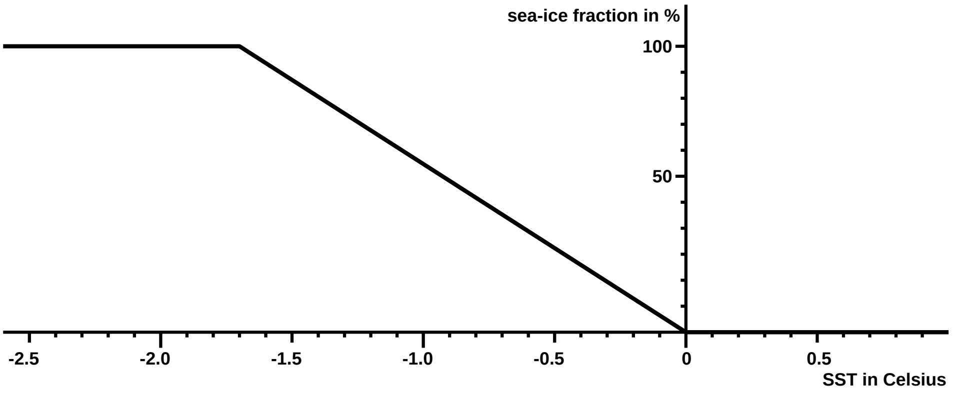
4.2.1 Low resolution data

Some of the input data from the former climate generation is still used in the new climate generation. It was transformed into HDF-format and stored in one HDF file, which includes the following data sets:

- usavn18
- alb_sea
- sea_ice (monthly data sets)
- rain (monthly data sets)
- soil_moist (monthly data sets)
- ts_ncar (monthly data sets)
- sst (monthly data sets)
- soil_eur

The data set names are mostly self descriptive. usavn18 includes elevation data from the *US Navy Tape*. This data was utilized to retrieve the orography in the former climate generation. It is now only necessary for the determination of the surface temperature from ts_ncar (section 4.4). The data set soil_eur includes the *FAO-Unesco Soil Database* (Bringfelt et al. 1995). Except of soil_eur, the data sets cover the whole globe in different coarse resolutions of a regular geographic grid. usavn18, rain, soil_moist and ts_ncar are given in 5° resolution. alb_sea and sst are in 1° resolution. Finally, soil_eur is in a 0.5° grid, but covers only Europe. Bringfelt (et al. 1995) and Bringfelt (1996) give a more detailed description for these data sources.

The sea_ice data set was deduced as monthly changing field from sst with the following simple function in a resolution of 1°:



HIRLAM uses this sea-ice data as a first guess. It is, however, desirable to include information about the salinity of the ocean in the future.

4.2.2 High resolution data

The preprocessed new high resolution physiographic data is stored in several HDF files containing data in different grid resolutions separately for each model domain of the operational HIRLAM setup. This would actually not be necessary, because the routine for retrieving the data sets in the climate generation is able to scan many different HDF files for finding a suitable data set. But a separation for different output resolutions is necessary for the calculation of the roughness length due to orography, because this property depends on the resolution of the HIRLAM grid (section 3). The distinction between the model domains also helps to save disk space and makes the climate generation faster by minimizing input file scans.

Table 3 shows the attributes of the high resolution input files, which are used in the respective climate generation of each of the operationally used HIRLAM model domains. All files contain similar data sets. They supply the new physiographic data in regular as well as in rotated geographical coordinates. The latter correspond to the respective rotated grid used in the model and data is supplied in two different resolutions. The first corresponds to the grid resolution of the respective operational model and is used for retrieving properties which depend on grid resolution like the sub grid orographic roughness length . The resolution in the second data file is higher and used for all other data. From the data given in regular coordinates the land–sea mask and the land–water mask are needed when processing data sets from the low resolution data.

Each file thus includes the following data sets:

- landseamask
- landmask
- height
- varnxn
- z0_oro_unsc
- cropland
- cropland_irrig
- grassland_short
- grassland_tall
- evergrn_needle
- evergrn_broadleaf
- decid_needle
- decid_broadleaf
- forest_mixed
- forest/fields
- field/woods
- evergrn_shrubs
- decid_shrubs
- tundra/wetland
- bogs
- ice
- desert
- semi-desert
- urban

Each HDF file contains apart from these data sets the two dimension data sets longitude and latitude. They contain information about the represented domain and the resolution of the connected data sets. Finally, general attributes like the information about the type of the coordinate system are stored as global attributes in the HDF file.

We take a closer look at the data sets now. The landseamask is a fraction field distinguishing between sea and continents. It does not contain inland water like lakes and rivers. These are represented in the landmask data set, which distinguishes between water in general and land. A field for the fraction of lakes can be deduced by taking the difference between these two data sets. The height field contains the unfiltered orography. The data set varnxn represents the variance of elevation height referred to the resolution of the region covered in the HDF file. The numbers nxn tell about the number of points used to determine the variance. This field can be used to determine a roughness length due to orography. But in in the actual setup of the climate generation, however, z0_oro_unsc is used (section 4.4). It contains the unscaled roughness length referred to the actual resolution in the HDF file and is determined from the variance (see section 3). All the following data sets represent fields of the fraction of occurrence for the respective land use class within a grid point. The HDF file only contains data sets of those land use classes, which occur in the represented region.

Supplying the high resolution input data already in the same rotated grid as the output diminishes the climate generation runtime and leads to a better representation of the variables. An example demonstrating the difference between a data retrieval from the regular grid and from the rotated grid is shown in Figure 3. The landmask was retrieved for a 0.45° rotated grid like it is used in HIRLAM-G, with an interpolation resolution of 0.05°. The input data resolution was 0.15° in both cases. Although the difference decreases with the use of higher interpolation resolutions, it is preferred to perform the transformation to the rotated grid in the highest possible resolution, i. e. within the physiographic preprocessing.

4.3 Parameters

There are two different kinds of parameter tables for the meteorological properties distinguishing between monthly changing properties and properties which are set constant in time. Furthermore, there are parameter tables which define general land surface types like a forest fraction, low vegetation fraction and no vegetation fraction (Table 7). These are determined from the land use classes and are utilized in the ISBA scheme.

One of the parameter tables for the monthly changing properties is shown in Table 5. The values for the leaf area index, roughness length and vegetation cover are taken from Bringfelt (1996), except for the land use classes forest/fields and field/woods, which don't occur in the tables of Bringfelt. Settings for these two classes were found with the help of the other land use classes taking the percentage of coverage with forest and open land into account (Table 4). Values for these percentages are defined in the parameter table shown in Table 7.

There are currently three parameters which are not supplied with monthly changing values. These are the albedo over land α, which is treated as a prognostic variable in HIRLAM, the depth of the total soil layer d and the minimum surface resistance r_{smin}. Values for them are also given by Bringfelt (1996). Table 6 shows the albedo parameter table as an example.

There is also one parameter file which describes the domain the climate generation is to be run for. It includes the following informations:

- southwestern coordinates of the domain in decimal degrees
- grid resolution in degrees (both coordinates)
- number of grid points (both coordinates)
- geographical position of the south pole (longitude, latitude) and angle of rotation

As an example, the region parameters for the current HIRLAM-D domain are:

```
-36.675 -15.177
0.05 0.05
182 170
80 0 0
```

4.4 Steps

The climate generation is performed in the following steps:

- retrieval of the land-sea mask from the high resolution elevation data.
- retrieval of the land mask (lakes, rivers etc. not part of the land fraction) from the high resolution land use data.
- filtering the land mask with a contrast filter.
- creation of the fraction of inland water.
- retrieval of elevation from the high resolution elevation data.
- filtering elevation with the shapiro filter.
- retrieval of the fractions of land use categories used in the ISBA scheme and defined in the parameter file shown in Table 7.
- retrieval of the monthly changing parameters (e.g. Table 5).
- retrieval of parameters not changing with month defined in the parameter files (e.g. 6).
- retrieval of FAO-Unesco soil classification types for Europe.
- retrieval of sst (interpolation to the required grid).
- retrieval of sea/sea-ice mask (interpolation to the required grid).
- retrieval of surface soil moisture (interpolation to the required grid).
- retrieval of usavn18, tsurf, rain, greenland-mask for the calculation of snow_depth, surface temperature and soil temperatures.
- calculation of snow_depth, surface temperature, soil temperatures and geopotential.
- retrieval of albedo over land.
- retrieval of albedo over sea from the ncar surface albedo.
- calculation of a combined albedo field taking sea-ice into account.
- retrieval of unscaled roughness due to orography and scaling.
- retrieval of roughness due to the land surface type.
- blending the roughness fields.

- filtering the blended roughness with a Gaussian filter.
- transformation into GRIB.

The data retrieval is connected with an interpolation, where the resolution of the input data is lower than the requested resolution. Otherwise, where the input resolution is higher than the requested one, an averaging method or other mathematical operation is applied, depending on the specification in the call of the retrieval routine (section 4.1).

The different land–water masks serve different filtering purposes in the climate generation process. In order to improve the coastline representation in the land fraction field, a contrast filter for sharpening the structures is applied. This is especially important over an area like Denmark (Figure 4).

For the calculation of the geopotential, the retrieved high resolution elevation data is used. It is filtered with a Shapiro filter (Shapiro 1970 and Huang 1996). The calculation of monthly mean surface temperatures and soil temperatures from ts_ncar is performed considering this surface height. But ts_ncar is referred to the heights of usavn18. This has to be taken into account when calculating the surface temperatures. $T_{surf,m}$ for month m is then determined from

$$T_{surf,m} = T_{ncar,m}\Gamma\left(h - h_{usavn18}\right), \tag{3}$$

where Γ denotes the dry adiabatic lapse. The mid-layer and the deep-layer monthly soil temperatures are determined from T_{surf} considering a phase lag in the soil layer(Bringfelt et al. 1995):

$$T_{mid,m} = 0.1\overline{T_{surf}} + 0.9\left(0.2T_{surf,m-1} + 0.8T_{surf,m}\right) \tag{4}$$

and

$$T_{deep,m} = 0.23\overline{T_{surf}} + 0.77\left(0.5T_{surf,m-1} + 0.5T_{surf,m}\right) \tag{5}$$

$\overline{T_{surf}}$ denotes the annual average of the surface temperature.

The albedo for sea and land is retrieved separately. While the albedo over sea is sufficiently represented in a coarse resolution of originally 1°, the albedo over land is determined from the land use classes. This is defined in the parameter table shown in Table 6.

For the calculation of the roughness fields the roughness length due to orography is determined from z0_oro_unsc with the following logarithmic scaling:

$$z_{0,oro} = a\left(z_{0,oro,unsc}\right)^B, \tag{6}$$

where

$$\begin{aligned} a &= 0.4038 \\ B &= 0.715. \end{aligned}$$

This scaling leads to roughness values of about 3 m over the mountains of Norway and about 5 m over the alps in the HIRLAM-G domain and the HIRLAM-E domain, which is in the magnitude of the former roughness fields. But further investigation may be necessary to refine this scaling. The roughness length decreases significantly over mountainous regions in the HIRLAM-D domain, because of the decreasing grid area used for determining $z_{0,oro,unsc}$.With growing resolution of the model grid, the influence of orography is more and more overtaken by the model dynamics and this is reflected in the determination process of $z_{0,oro,unsc}$ (section 3). Figures 5 – 7 show the complete roughness fields over the southern Norwegian mountains for the different model domains as an example for this change. The roughness fields of HIRLAM-G and HIRLAM-E are dominated by z_0,oro, but in the HIRLAM-D domain, z_0,oro is not that dominant any more and structures showing the influence of vegetation can be recognized.

The vegetation roughness is retrieved from the land use information as a monthly changing parameter. The combination of both roughness lengths is performed according to Tibaldi and Geleyn (1981):

$$z_0 = \sqrt{z_{0,oro}^2 + z_{0,veg}^2}. \tag{7}$$

This blending is done by taking the landmask into account with a fraction limit of 5%. Finally, a Gaussian filter is applied to the blended roughness length.

4.5 Generated fields

Table 8 shows a list of all variables in the output GRIB file. The number of parameters may still change with further future developments of the climate generation.

The most significant changes compared to the former climate fields occur in the surface geopotential, fraction of land and roughness length. To give an impression of the resolution and quality of the fields, these parameters are shown in Figures 8 to 10 for the region of Denmark in two different model representations.

The surface geopotential and roughness length over Greenland are shown in Figure 12. The roughness length over ice is set to 0.01 m according to Bringfelt (1996). Over the coastal areas of Greenland, roughness is almost completely determined by sub grid orographic variance, which leads to high values in there. As most of the synoptic observation stations in Greenland are located near the coast within the areas of highly variable orography, it is interesting to see, how the elevations of them are represented in the model data. Table 9 lists the elevations for the HIRLAM-G and HIRLAM-N domain and gives an impression of the improved representation when using the new physiographic data.

The new climate files already contain fields which are to be utilized in the ISBA scheme for HIRLAM. Figure 11 shows the vegetation fraction for two months for the domain of HIRLAM-E. New included are also the fraction of barren land (no vegetation), fraction of forest and the soil type data. The latter currently covers only Europe.

4.6 Comments

The new climate generation can be set up in different ways and allows it to apply many different procedures on the input data like interpolation, different kinds of averaging, scaling or mask operations. In order to establish a well working climate generation with a reasonable output it is necessary to be aware about the behaviour of the programs involved in the data processing. As almost all procedures which are mentioned above are performed by the routine `ctopo`, this is the main point of focus.

An exceptional behaviour of routine `ctopo` is worth mentioning here. It is convenient with `ctopo` to retrieve and calculate field properties using a landmask. But care has to be taken when doing this, because this can lead to unwelcome smooth field patterns over coastal areas. An example which demonstrates the difference between utilizing a landmask and using another approach is given for the roughness length as follows.

We regard the roughness length derived from the land use classes (z0_veg). The first possible way to retrieve this variable is to utilize the landmask when creating the field with `ctopo`. It should then not be necessary to set values for water in the parameter file defining z0_veg. Figure 13 shows the resulting roughness field from this retrieval over Denmark as represented in HIRLAM-E. The representation at the coasts is rather disappointing. If we now introduce the parameter setting for the roughness length over water in the parameter file (see for example Table 5) and still use the landmask when retrieving the data with `ctopo`, the field as shown in Figure 14 is retrieved. It gives a significantly better representation of the coastal areas. But it should actually not be necessary to use both the landmask and the parameter setting for water roughness. So if we now determine z0_veg without using the landmask in the call to `ctopo` but only utilize the parameter setting, the coast representation becomes even better (Figure 15).

This difference between using and not using the landmask during the data retrieval of z0_veg does also occur in the other HIRLAM domains. It is bigger in the HIRLAM-G resolution and smaller in the resolution of HIRLAM-D. But even in the latter the difference is quite important, as the Figures 16 – 18 show. They refer to the respective calculation processes which are described above.

The question whether to use a landmask at data retrieval or not is also important when retrieving and scaling the roughness length due to sub grid orography z0_oro. Again, the coastal representation becomes better when no landmask is used.

The utilization of a landmask during the blending process of z0_oro and z0_veg, however, does not have such negative effects as described above.

5 Closing Remarks

New high resolution physiographic data were included in a new climate generation for the HIRLAM system. They reveal much more details in the deduced parameter fields than the old data bases. Although the data bases cover the whole globe and are given in a resolution of about 1 km, the precision of the land use data was not sufficient over some areas of Denmark. It had to be improved with the additional inclusion of local data from Denmark in order to suit the need for precise details in the HIRLAM-D domain.

The new software for the climate generation simplifies some technical aspects of the climate generation like the use of the universal file format HDF. A physiographic preprocessing step is performed before the climate generation in order to supply the input data and to reduce extreme computer resource demands during the climate generation. However, the new climate generation still needs more resources than the old one, because it deals with fields of higher resolution and creates additional fields. The concept of parameter tables within the climate generation allows to add properties, which are determined from the land use data, easily by including them in the parameter tables.

The new climate generation is to be included into the HIRLAM reference system in the near future. There are still issues to be further developed like the calculation of the sub grid orographic roughness or the sea ice determination. New parameters may be added in the future depending on the needs for the HIRLAM simulations.

Acknowledgements

The code to the new climate generation was provided by Han The from KNMI. He also wrote a front-end library to HDF for easier use of HDF and gave some support. Basic routines and information for accessing the high resolution topographic data bases were provided by Jens Hesselbjerg Christensen and Ole Bøssing Christensen. I would further like to thank Jess U. Jørgensen for supplying the interface between ASIMOV and DMI–GRIB. He also checked some of the generated files. Bjarne Amstrup provided the verification of the orography for the Greenland model with the actual station heights.

References

Bringfelt B., Gustafson N., Vilmusenaho P., Järvenoja S. 1995: Updating the HIRLAM physiography and climate data base. *HIRLAM Tech. Rep.*, **19**.

Bringfelt B., 1996: Tests of a new Land Surface Treatment in HIRLAM. *HIRLAM Tech. Rep.*, **23**.

Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J., and Wilson, M. F., 1986: Biosphere-atmosphere transfer scheme (BATS) for the NCAR community climate model. *NCAR Technical Note*, **NCAR/TN275+STR**, Boulder, CO. 69 p.

The H., 1998: Revised routines for pre-processing the HIRLAM physiographic data sets (script Preps). *KNMI internal note*, 11p.

NCSA 1998: HDF User's Guide, Version 4.1r2. *NCSA Scientific Data Technologies*, http://hdf.ncsa-.uiuc.edu. 232p.

Olson, J. S., 1994a: Global ecosystem framework-definitions. *USGS EROS Data Center Internal Report*, Sioux Falls, SD, 37 p.

_ 1994b: Global ecosystem framework-translation strategy. *USGS EROS Data Center Internal Report*, Sioux Falls, SD, 39 p.

Sass, B. H., 1998: Operational HIRLAM at DMI. *HIRLAM Newsletter*, **31**, 57–59.

Shapiro, R. 1970: Smoothing, Filtering and Boundary Effects. *Rev. Geophys. and Space Phys.*, **8**, 359–387.

Steinwand, D. R., 1994: Mapping raster imagery to the Interrupted Goode Homolosine projection *International Journal of Remote Sensing*, **15**, No. 17, 3,463–3,472.

Tibaldi, S., Geleyn, J.–F., 1981: The Production of a new Orography, Land-Sea Mask and associated Climatological Surface Fields for Operational Purposes. *ECMWF Research Department Technical Memorandum*, **40**, 13 p. + Appendices and Figures.

USGS 1997: The Global Land Cover Characteristics Data Base. *U.S. Geological Survey (USGS)*, http://edcwww.cr.usgs.gov/landdaac/glcc/glcc.html.

USGS 1998: GTOPO30, Global 30 Arc Second Elevation Data Set. *U.S. Geological Survey (USGS)*, http://edcwww.cr.usgs.gov/landdaac/gtopo30/gtopo30.html

Huang, X.-Y. 1996: Horizontal Diffusion and Filtering in a Mesoscale Numerical Weather Prediction Model. *DMI Scientific report*, **96-5**.

Nr	Global Ecosystems Legend	used classification
1	Urban	urban
2	Low Sparse Grassland	grassland short
3	Coniferous Forest	evergreen needle
4	Deciduous Conifer Forest	deciduous needle
5	Deciduous Broadleaf Forest	deciduous broadleaf
6	Evergreen Broadleaf Forests	evergreen broadleaf
7	Tall Grasses and Shrubs	grassland tall
8	Bare Desert	desert
9	Upland Tundra	tundra/wetland
10	Irrigated Grassland	cropland irrigated
11	Semi Desert	semi-desert
12	Glacier Ice	ice
13	Wooded Wet Swamp	bogs
14	Inland Water	water
15	Sea Water	ocean
16	Shrub Evergreen	evergreen shrubs
17	Shrub Deciduous	deciduous shrubs
18	Mixed Forest and Field	forest mixed
19	Evergreen Forest and Fields	forest/fields
20	Cool Rain Forest	evergreen needle
21	Conifer Boreal Forest	evergreen needle
22	Cool Conifer Forest	evergreen needle
23	Cool Mixed Forest	forest mixed
24	Mixed Forest	forest mixed
25	Cool Broadleaf Forest	deciduous broadleaf
26	Deciduous Broadleaf Forest	deciduous broadleaf
27	Conifer Forest	evergreen needle
28	Montane Tropical Forests	evergreen broadleaf
29	Seasonal Tropical Forest	evergreen broadleaf
30	Cool Crops and Towns	cropland
31	Crops and Town	cropland
32	Dry Tropical Woods	deciduous broadleaf
33	Tropical Rainforest	evergreen broadleaf
34	Tropical Degraded Forest	evergreen broadleaf
35	Corn and Beans Cropland	cropland
36	Rice Paddy and Field	cropland irrigated
37	Hot Irrigated Cropland	cropland irrigated
38	Cool Irrigated Cropland	cropland irrigated
39	Cold Irrigated Cropland	cropland irrigated
40	Cool Grasses and Shrubs	grassland short
41	Hot and Mild Grasses and Shrubs	grassland short
42	Cold Grassland	grassland short
43	Savanna (Woods)	grassland tall
44	Mire Bog Fen	bogs
45	Marsh Wetland	bogs
46	Mediterranean Scrub	evergreen shrubs
47	Dry Woody Scrub	evergreen shrubs
48	Dry Evergreen Woods	evergreen broadleaf
49	Volcanic Rock	desert
50	Sand Desert	desert

Nr	Global Ecosystems Legend	used classification
51	Semi Desert Shrubs	semi-desert
52	Semi Desert Sage	semi-desert
53	Barren Tundra	tundra/wetland
54	Cool S Hemisphere Mixed Forests	forest mixed
55	Cool Fields and Woods	field/woods
56	Forest and Field	forest/fields
57	Cool Forest and Field	forest/fields
58	Fields and Woody Savanna	field/woods
59	Succulent and Thorn Scrub	deciduous shrubs
60	Leaf Mixed Woods	deciduous broadleaf
61	Deciduous Mixed Boreal Forest	forest mixed
62	Narrow Conifers	evergreen needle
63	Wooded Tundra	tundra/wetland
64	Heath Scrub	evergreen shrubs
65	Coastal Wetland NW	bogs
66	Coastal Wetland NE	bogs
67	Coastal Wetland SE	bogs
68	Coastal Wetland SW	bogs
69	Polar and Alpine Desert	desert
70	Glacier Rock	desert
71	Salt Playas	desert
72	Mangrove	bogs
73	Water and Island Fringe	bogs
74	Land Water and Shore	bogs
75	Land and Water Rivers	bogs
76	Crop and Water Mixtures	bogs
77	S Hemisphere Conifers	evergreen needle
78	S Hemisphere Mixed Forest	forest mixed
79	Wet Sclerophylic Forest	forest mixed
80	Coastline Fringe	bogs
81	Beaches and Dunes	semi-desert
82	Sparse Dunes and Ridges	semi-desert
83	Bare Coastal Dunes	desert
84	Residual Dunes and Beaches	semi-desert
85	Compound Coastlines	bogs
86	Rocky Cliffs and Slopes	desert
87	Sandy Grassland and Shrubs	grassland tall
88	Bamboo	deciduous shrubs
89	Moist Eucalyptus	evergreen broadleaf
90	Rain Green Tropical Forest	deciduous broadleaf
91	Woody Savanna	field/woods
92	Broadleaf Crops	cropland
93	Grass Crops	cropland
94	Crops Grass Shrubs	cropland

Table 1: Transformation of the *Global Ecosystems Legend* to the BATS-like classification scheme used in HIRLAM.

new routines	former routine
clpp	clprep
ctopo	clifil, inmania
	pcsmder, m_soil, m_icesea
tsoil	pcn_ts, pcsnow, pcstderv, pcsst[1] pctsbln
albedo	m_albedo
z0merge	pcz0bln
sdsep	–
sharpen	–
smooth_gauss	pcgausf
smooth_sha	(shafil)

Table 2: The routines of the new climate generation and how they replace the present routines. Note that smooth_sha is just an interface to shafil.

	HIRLAM-G	HIRLAM-N	HIRLAM-E	HIRLAM-D
1. file				
grid	rotated	rotated	rotated	rotated
domain lon	-81.725° – 45.175°	-38.075° – 9.025°	-63.275° – -4.475°	-44.675° – -19.575°
domain lat	-55.527° – 65.973°	-14.277° – 35.223°	-37.677° – 22.623°	-23.177° – 1.323°
resolution	0.45°	0.15°	0.15°	0.05°
2. file				
grid	rotated	rotated	rotated	rotated
domain lon	-81.725° – 45.175°	-38.075° – 9.025°	-63.275° – -4.475°	-44.675° – -19.575°
domain lat	-55.527° – 65.973°	-14.277° – 35.223°	-37.677° – 22.623°	-23.177° – 1.323°
resolution	0.15°	0.05°	0.05°	0.0167°
3. file				
grid	regular	regular	regular	regular
domain lon	-180° – 180°	-100° – 80°	-56° – 58°	-5° – 25°
domain lat	13.5° – 90°	49.5° – 90°	30° – 78°	45° – 65°
resolution	0.45°	0.15°	0.15°	0.05°

Table 3: HDF input files for the climate generation separated for each model domain of the operational setup. A further description is given in section 4.2.

	forest/fields			field/woods		
month	LAI	z_0 in m	veg	LAI	z_0 in m	veg
01	2.0	0.150	0.60	1.7	0.051	0.45
02	2.1	0.154	0.61	1.8	0.056	0.47
03	2.3	0.163	0.63	2.0	0.061	0.50
04	2.9	0.186	0.65	2.4	0.064	0.55
05	3.5	0.210	0.69	2.9	0.080	0.60
06	3.8	0.230	0.72	3.4	0.098	0.65
07	4.5	0.250	0.74	3.8	0.100	0.69
08	4.2	0.242	0.73	3.7	0.099	0.66
09	3.6	0.220	0.72	3.4	0.084	0.61
10	3.0	0.201	0.67	3.0	0.065	0.57
11	2.1	0.161	0.64	2.5	0.060	0.50
12	2.0	0.152	0.62	2.0	0.056	0.47

Table 4: Parameter settings of leaf area index LAI, roughness length z_0 and vegetation cover veg for the two land use classes forest/fields and field/woods, which are not defined in Bringfelt (1996).

LAI, z0 and vegetation index Jan

	LAI_01	z0_veg_01	vegindx_01
evergrn_needle	5	1	0.8
decid_needle	0.5	0.812	0.65
decid_broadleaf	0.5	0.667	0.75
evergrn_broadleaf	6	2	0.99
forest_mixed	3	0.69	0.75
forest/fields	2	0.15	0.6
cropland	1.5	0.018	0.15
grassland_short	0.5	0.012	0.3
grassland_tall	0.5	0.051	0.45
tundra/wetland	0.5	0.02	0.3
cropland_irrig	1.5	0.018	0.15
semi-dessert	0.5	0.04	0.04
bogs	3	0.03	0.5
evergrn_shrubs	4	0.1	0.75
decid_shrubs	0.7	0.04	0.3
field/woods	1.7	0.051	0.45
desert	0	0.05	0
ice	0	0.01	0
urban	0	2	0
water	0	0.001	0

Table 5: Parameter table to convert the land use classes into monthly values of leaf area index LAI, vegetation roughness z0_veg and a vegetation fraction vegindx. The shown list is valid for January.

	albedo
# Forest + forest_interrupted	
evergrn_needle	0.16
decid_needle	0.16
decid_broadleaf	0.16
evergrn_broadleaf	0.14
forest_mixed	0.16
forest/fields	0.17
# Low vegetation	
cropland	0.18
grassland_short	0.20
grassland_tall	0.20
tundra/wetland	0.18
cropland_irrig	0.18
semi-dessert	0.20
bogs	0.14
evergrn_shrubs	0.18
decid_shrubs	0.17
field/woods	0.18
# no vegetation including urban	
dessert	0.48
ice	0.60
urban	0.20

Table 6: Parameter table to determine albedo values from the land use information.

	forest	low_veg	no_veg
evergrn_needle	1	0	0
decid_needle	1	0	0
decid_broadleaf	1	0	0
evergrn_broadleaf	1	0	0
forest_mixed	1	0	0
# Low vegetation			
cropland	0	1	0
grassland_short	0	1	0
grassland_tall	0	1	0
tundra/wetland	0	1	0
cropland_irrig	0	1	0
semi-desert	0	1	0
bogs	0	1	0
evergrn_shrubs	0	1	0
decid_shrubs	0	1	0
field/woods	0.05	0.95	0
forest/fields	0.25	0.75	0
# no vegetation including urban			
desert	0	0	1
ice	0	0	1
urban	0	0	1

Table 7: Parameter table which defines surface types used in the ISBA scheme.

parameter	level type	level	description
6	105	0	surface geopotential
84	105	0	surface albedo
83	105	0	surface roughness
11	102	0	sea surface temperature
11	105	0	surface temperature
11	105	999	deep soil temperature
11	105	998	climatic deep soil temperature
86	105	0	soil water
86	105	999	deep soil water
86	105	998	climatic deep soil water
66	105	0	snow depth
195	105	0	soil type
81	105	0	fraction of land
91	102	0	fraction of ice
79	105	0	fraction of urbanization
182	105	0	fraction of no vegetation land
87	105	0	fraction of vegetation
196	105	0	fraction of lakes
197	105	0	fraction of forest

Table 8: The parameters in the output GRIB file produced in the climate generation. Included are already fields utilized in the ISBA scheme.

		HIRLAM-G				HIRLAM-N			
		value		difference		value		difference	
station	height	old	new	old	new	old	new	old	new
04201 Qaanaaq	16	314	286	298	270	0	0	-16	-16
04202 Pituffik	77	577	538	500	461	594	274	517	197
04203 Kitsissut	11	8	4	-3	-7	3	5	-8	-6
04207 Hall Land	105	416	207	311	102	364	213	259	108
04208 Kitsissorsuit	40	91	37	51	-3	6	12	-34	-28
04210 Upernavik	120	218	121	98	1	167	74	47	-46
04214 Nuussuaq	27	289	173	262	146	173	114	146	87
04220 Aasiaat	40	44	4	4	-36	43	20	3	-20
04221 Ilulissat Lufthavn	29	253	201	224	172	101	28	72	-1
04228 Attu	12	49	38	37	26	6	3	-6	-9
04230 Sisimiut	12	176	172	164	160	138	51	126	39
04231 Kangerlussuaq	50	373	366	323	316	279	244	229	194
04242 Sioralik	14	96	112	82	98	38	14	24	0
04250 Nuuk	80	169	125	89	45	171	84	91	4
04253 Ukiivik	22	193	118	171	96	83	17	61	-5
04260 Paamiut	13	310	223	297	210	218	24	205	11
04266 Nunarsuit	33	117	69	84	36	11	1	-22	-32
04270 Narsarsuaq Lufthavn	8	782	932	774	924	489	277	481	269
04272 Qaqortoq	32	320	252	288	220	247	24	215	-8
04285 Angisoq	16	201	94	185	78	170	10	154	-6
04301 Kap Morris Jesup	4	387	69	383	65	596	106	592	102
04312 Station Nord AWS	34	153	-8	119	-42	355	6	321	-28
04313 Henrik Krøyer Holme	10	5	0	-5	-10	1	5	-9	-5
04320 Danmarkshavn	14	110	29	96	15	126	10	112	-4
04330 Daneborg	44	273	302	229	258	137	163	93	119
04339 Ittoqqortoormiit	68	125	100	57	32	173	160	105	92
04351 Aputiteeq	13	235	229	222	216	79	13	66	0
04360 Tasiilaq	50	557	532	507	482	254	137	204	87
04373 Ikermit	85	311	251	226	166	123	74	38	-11
04382 Ikermiuarsuk	39	323	190	284	151	203	34	164	-5
04390 Prins Christian Sund	88	335	44	247	-44	430	66	342	-22

Table 9: Verification of the elevation heights of the synoptic observation stations over Greenland for the model domains of HIRLAM-G and HIRLAM-N. Shown are both the values from the old climate files and the new data. Some of the differences between model representation and station height still remain due to the smoothing of the orography field with a filter during the climate generation (section 4.4).

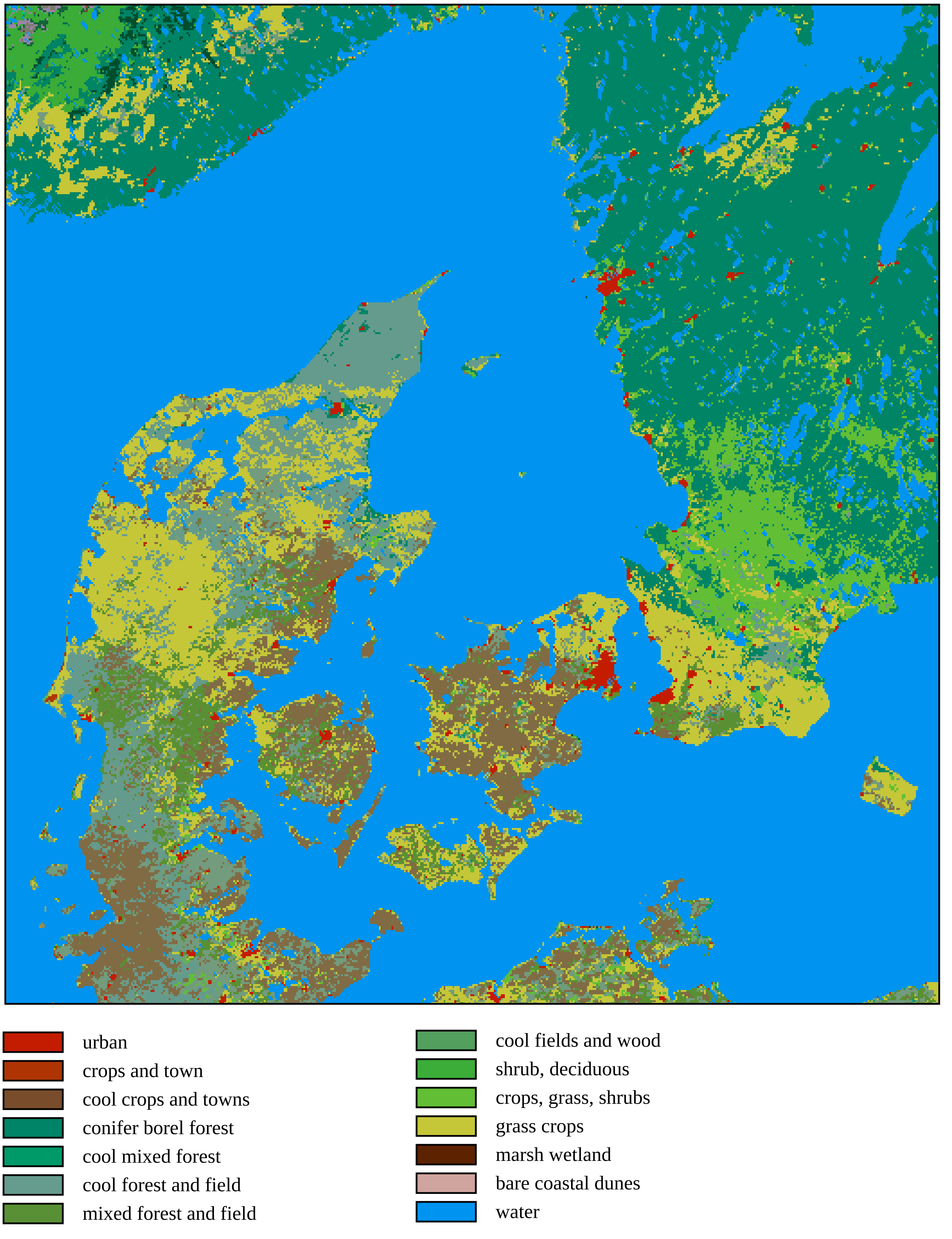


Figure 1: Land use over Denmark in the *Global Ecosystems Legend* from the GLCC-data in an equal area projection. For convenience, only the land use classes important for this region are listed in the legend.

Figure 2: Land use over Denmark in the *Global Ecosystems Legend* including the KMS-data in an equal area projection. For convenience, only the land use classes important for this region are listed in the legend.

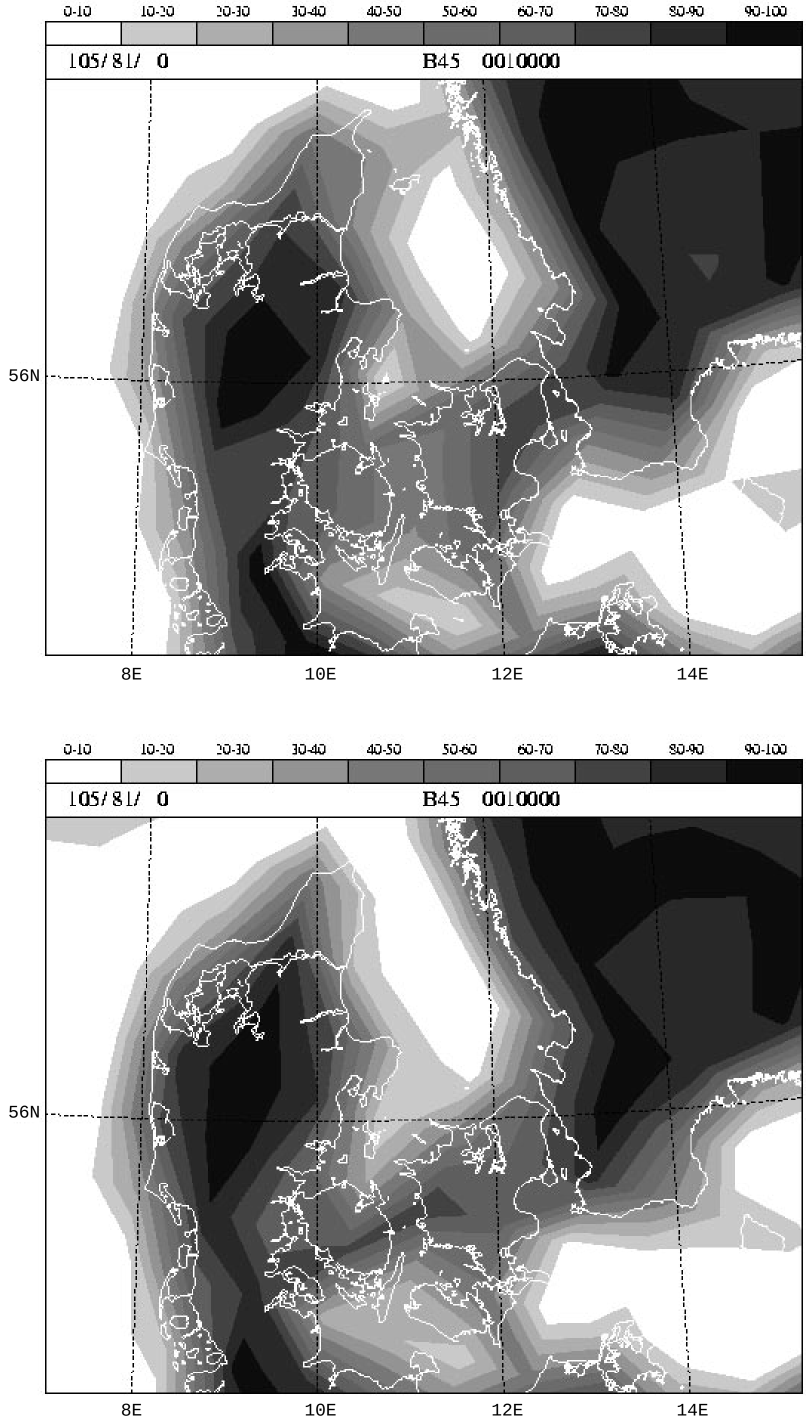
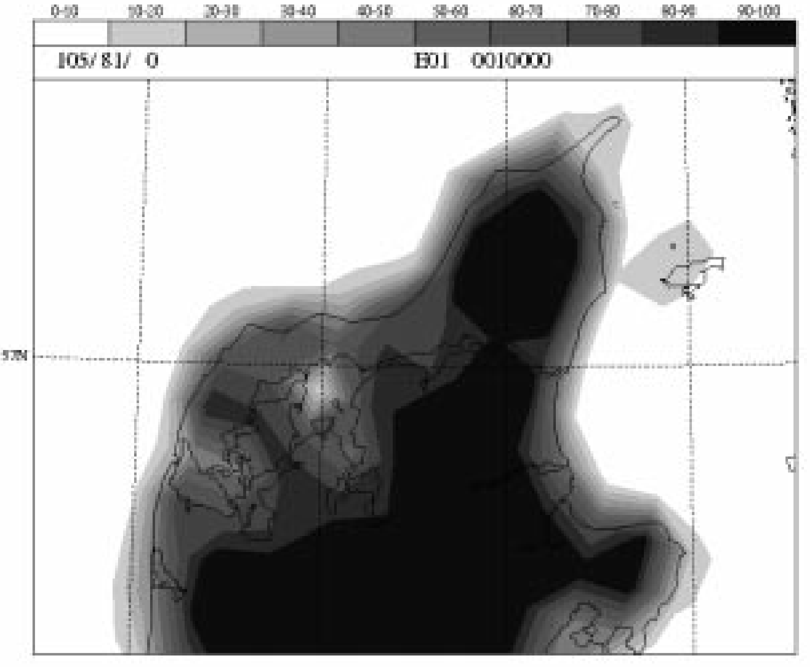


Figure 3: Comparison between the land mask retrieved from a regular grid (upper) and the land mask retrieved from the rotated grid with the same rotation parameter setting as the output grid (lower). In both cases the resolution of the input data was three times higher than the resolution of the output grid (0.45°).



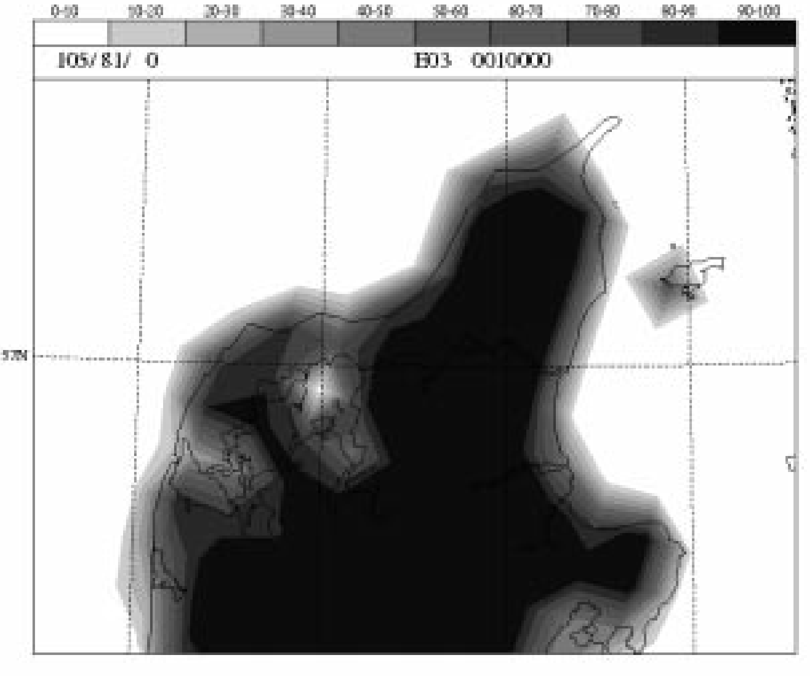


Figure 4: Comparison of the fraction of land for North Jutland as it is represented in HIRLAM-E, without a contrast filter (upper) and with contrast filter (lower).

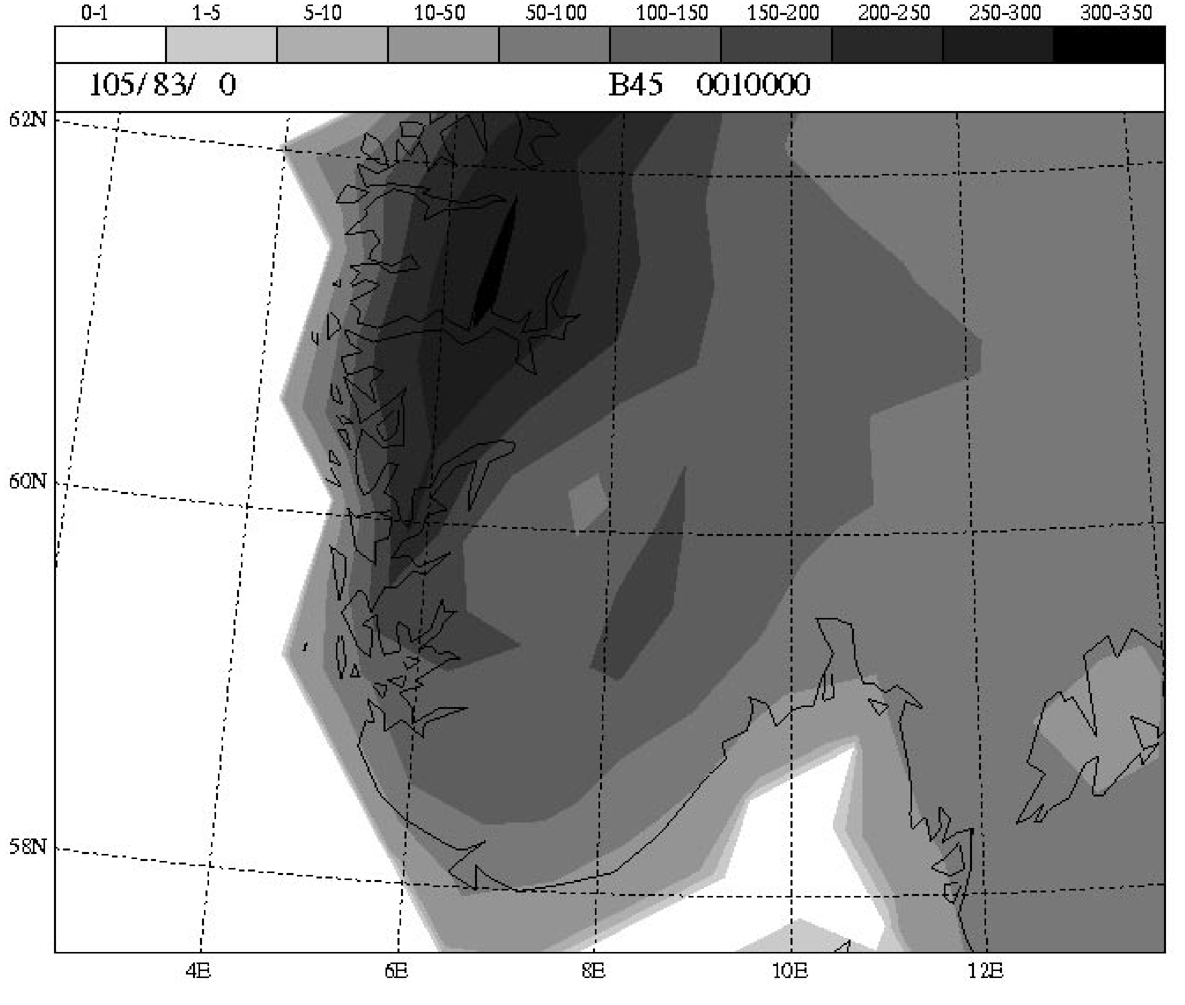


Figure 5: Roughness length in cm for the southern part of Norway and how it is represented in HIRLAM-G. In this area roughness length is mainly influenced by sub grid orography.

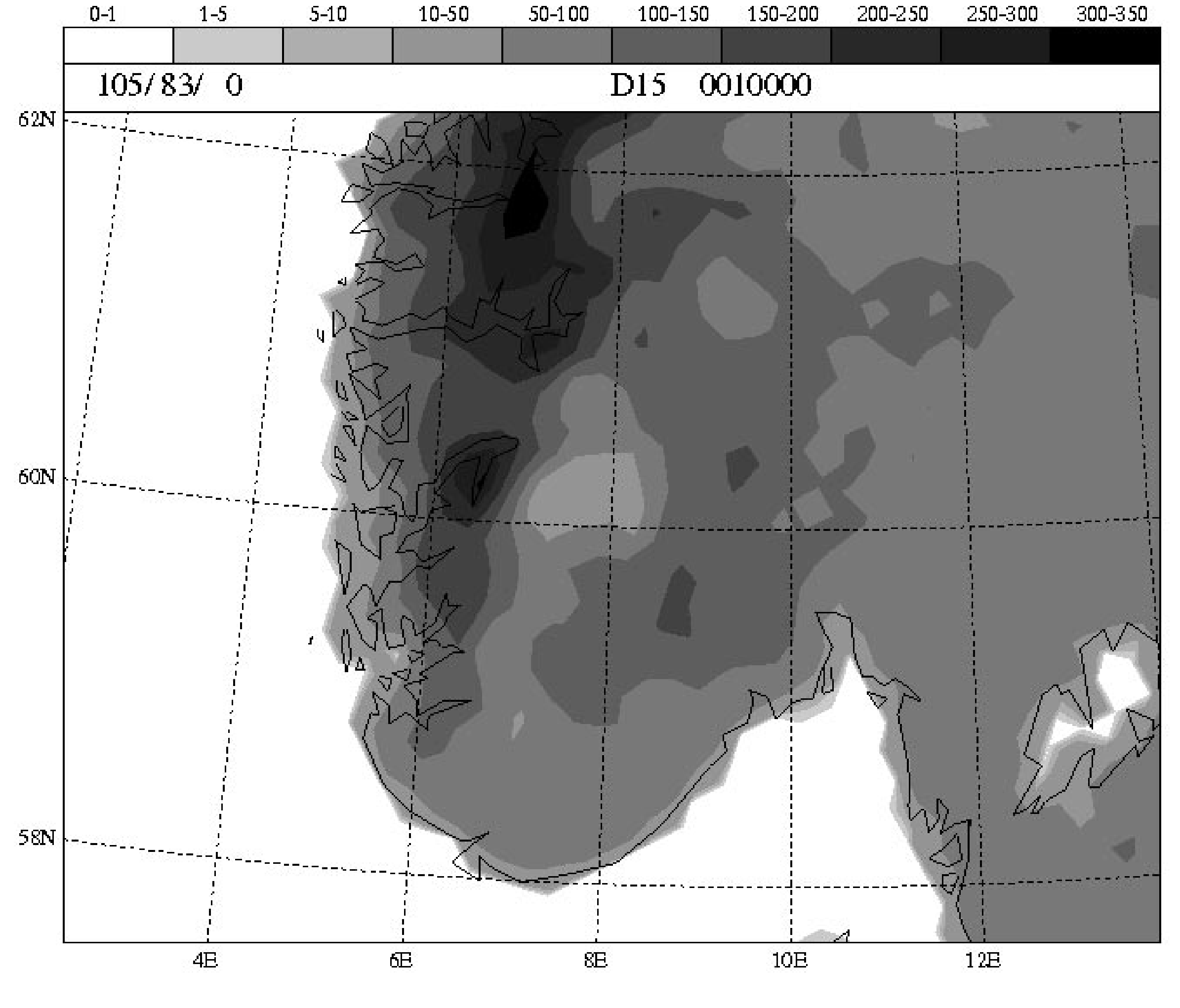


Figure 6: Roughness length in cm for the southern part of Norway and how it is represented HIRLAM-E. The roughness length is in the mountainous region is still mainly influenced by sub grid orography.

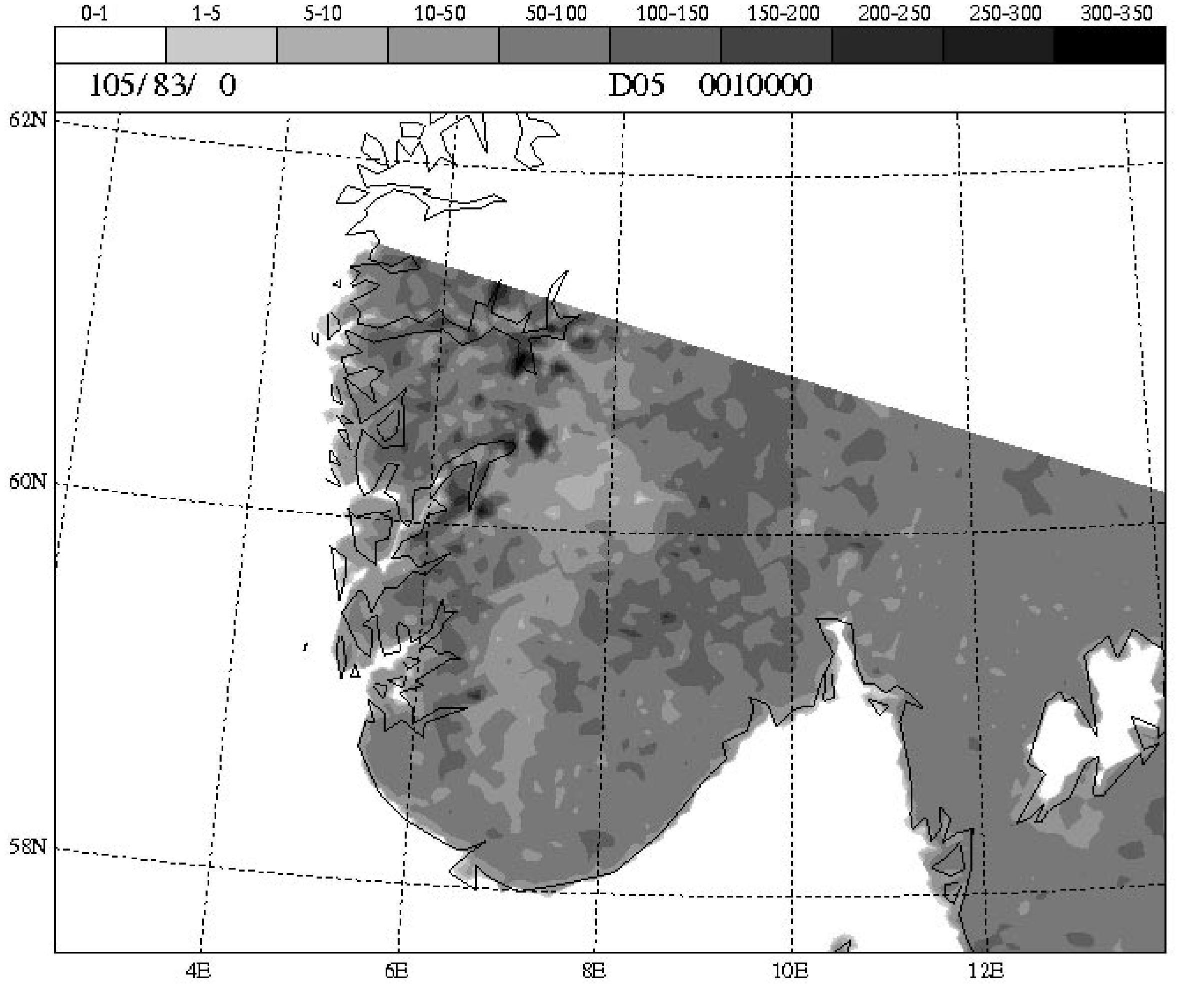


Figure 7: Roughness length in cm for the southern part of Norway and how it is represented in HIRLAM-D. The dominance of the roughness length due to sub grid orography over the mountains decreases with increasing resolution of the model grid. The model domain of HIRLAM-D ends within the shown area (white area in the upper part of the picture).

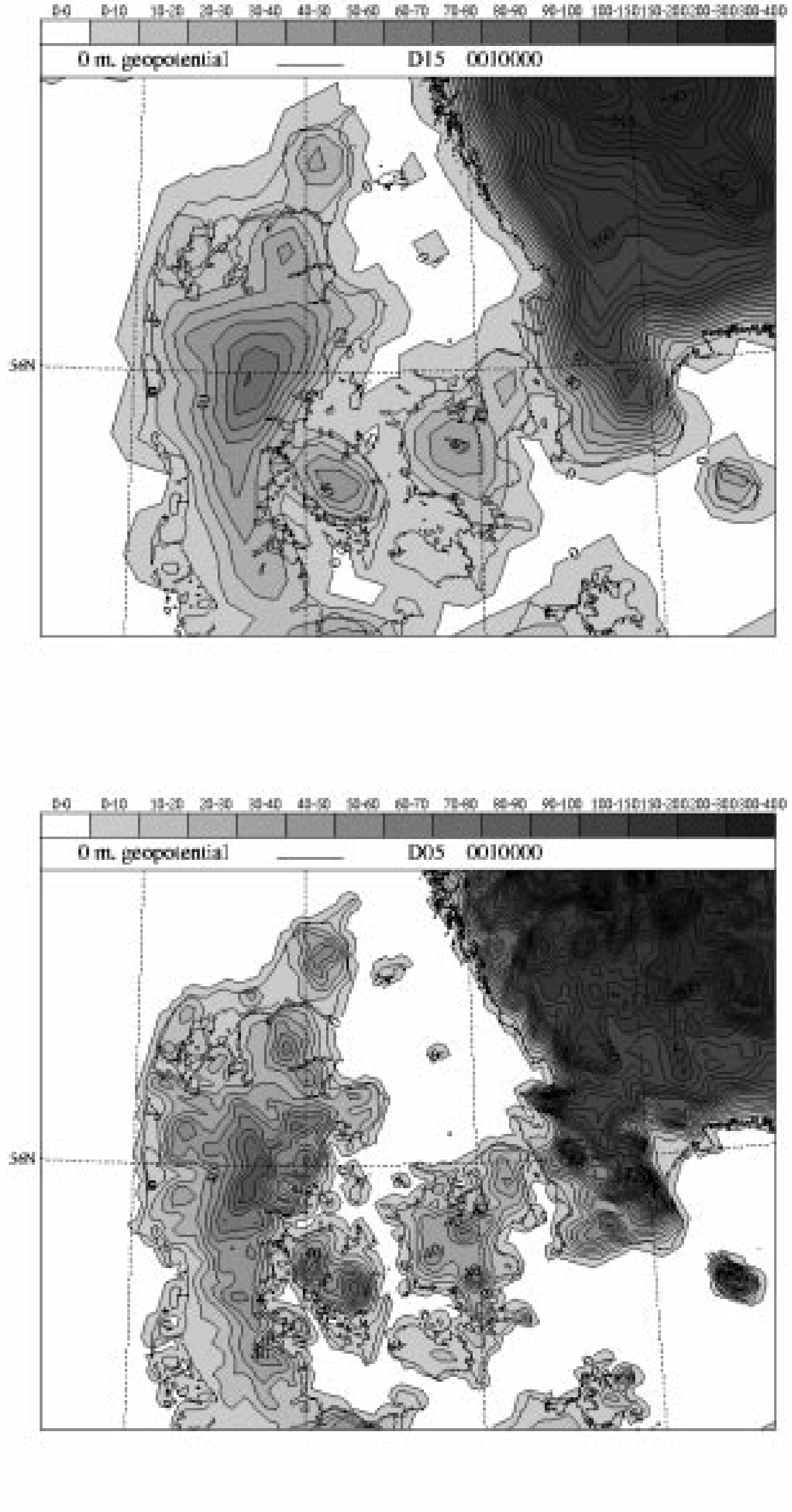
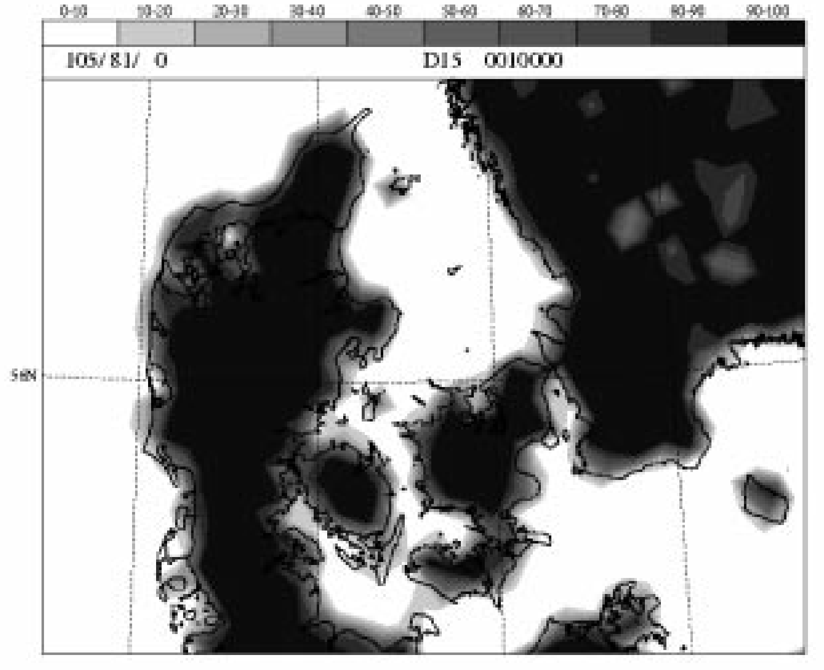


Figure 8: Surface geopotential in m retrieved from the high resolution data base and its representation in HIRLAM-E (upper) and HIRLAM-D (lower). The contour interval is 10 m.



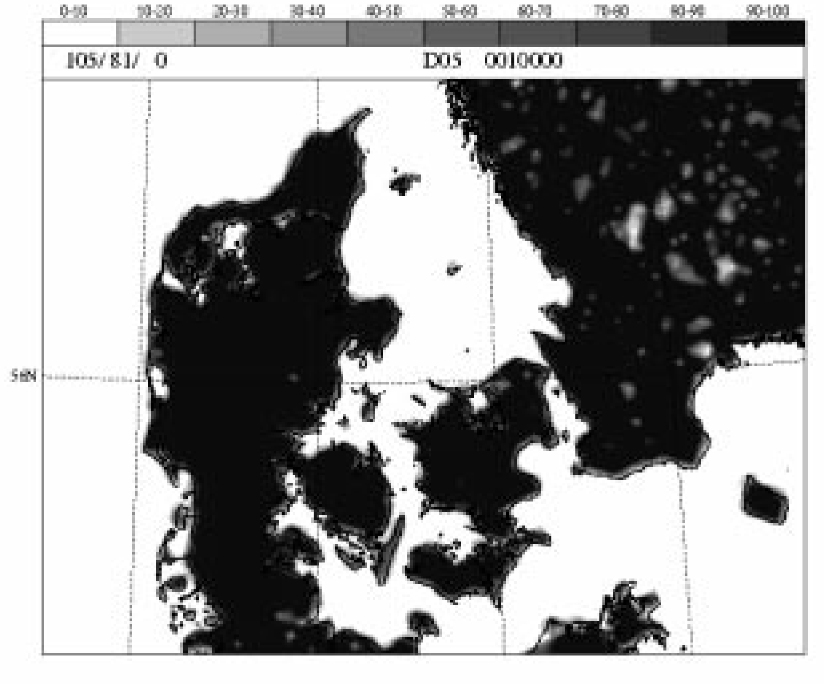


Figure 9: Fraction of land over Denmark in % for HIRLAM-E (upper) and for HIRLAM-D (lower).

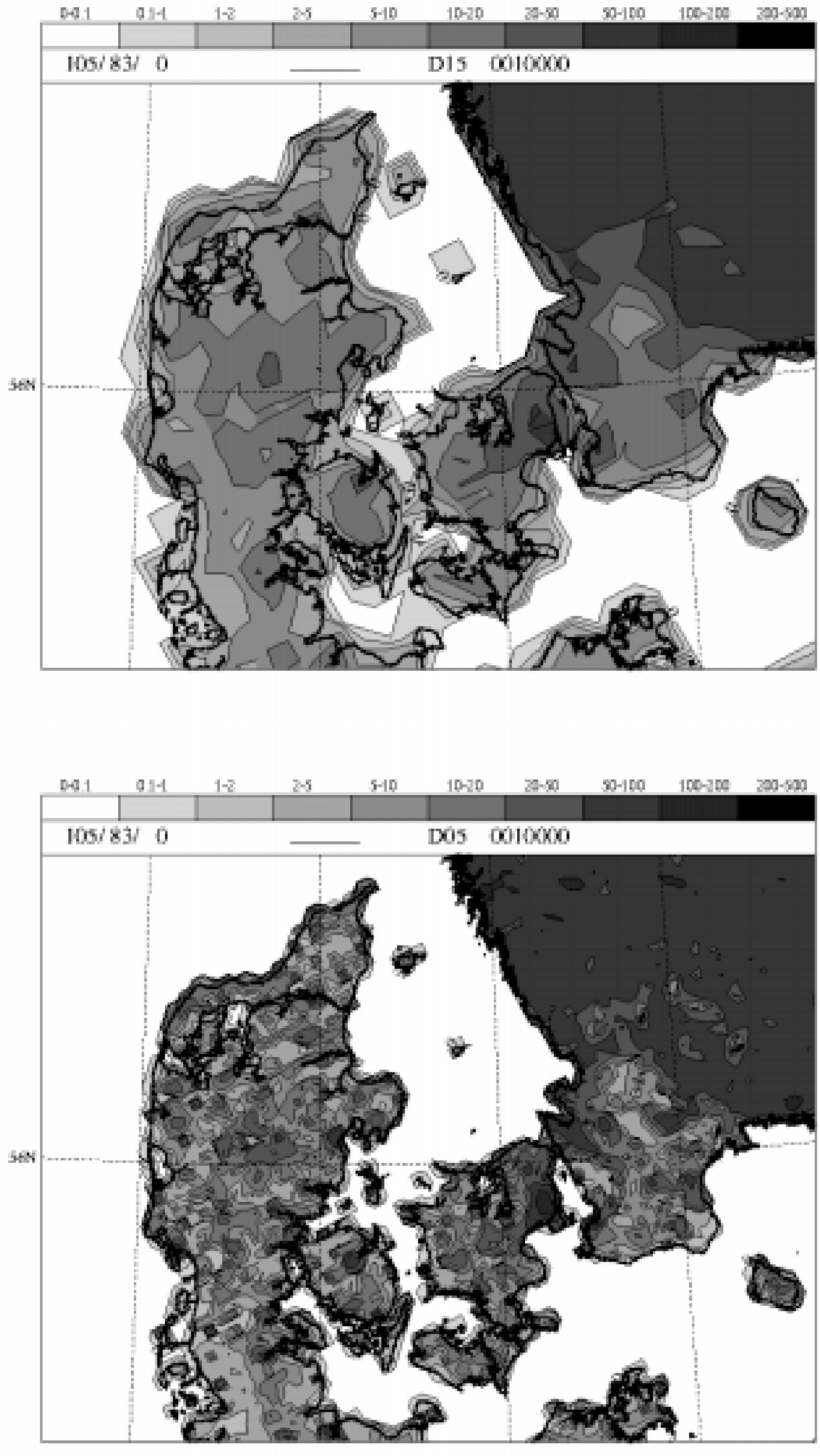


Figure 10: Roughness length in cm for January as it is represented in HIRLAM-E (upper) and HIRLAM-D (lower).

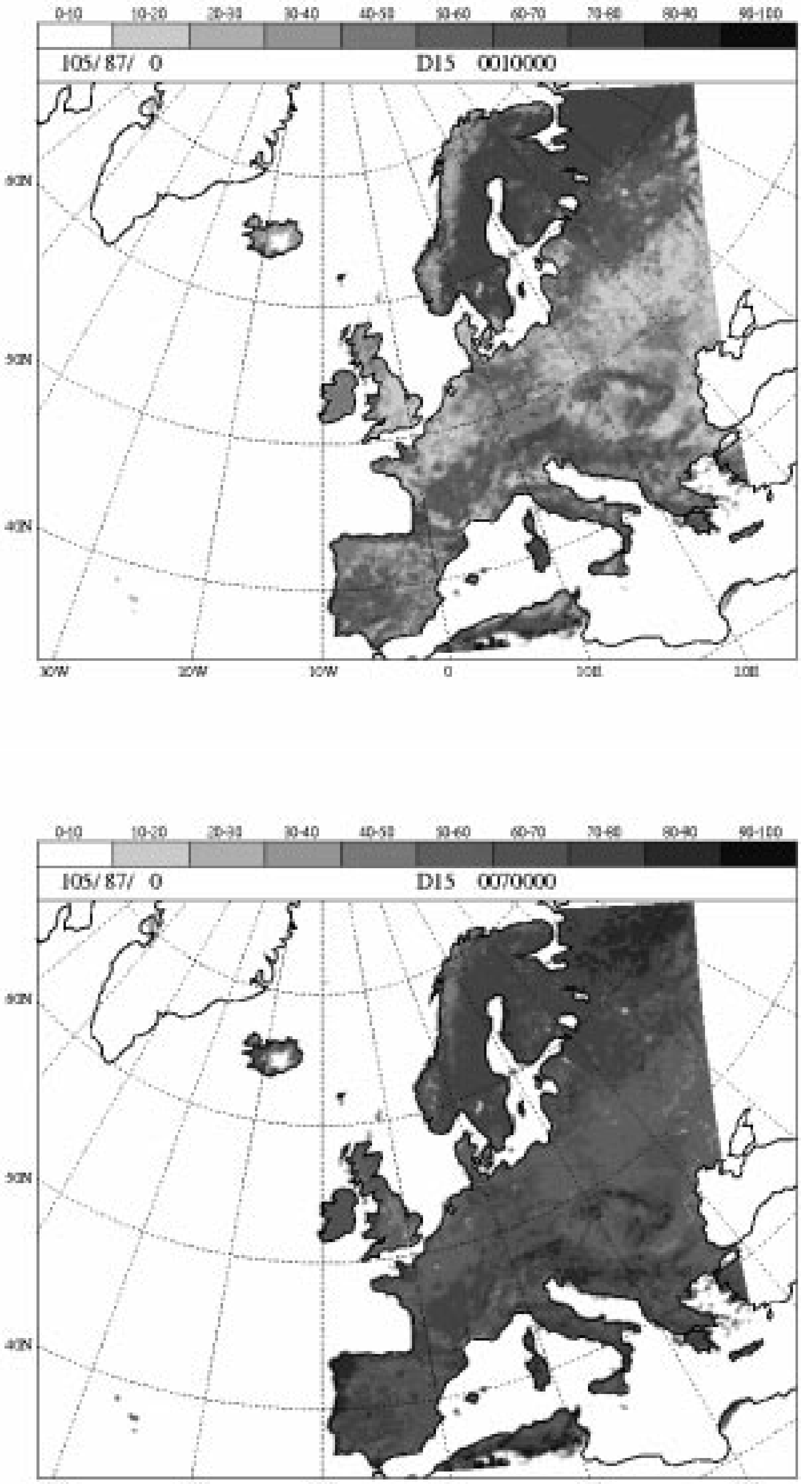


Figure 11: Fraction of vegetation in % for January (upper) and July (lower) for the domain of HIRLAM-E.

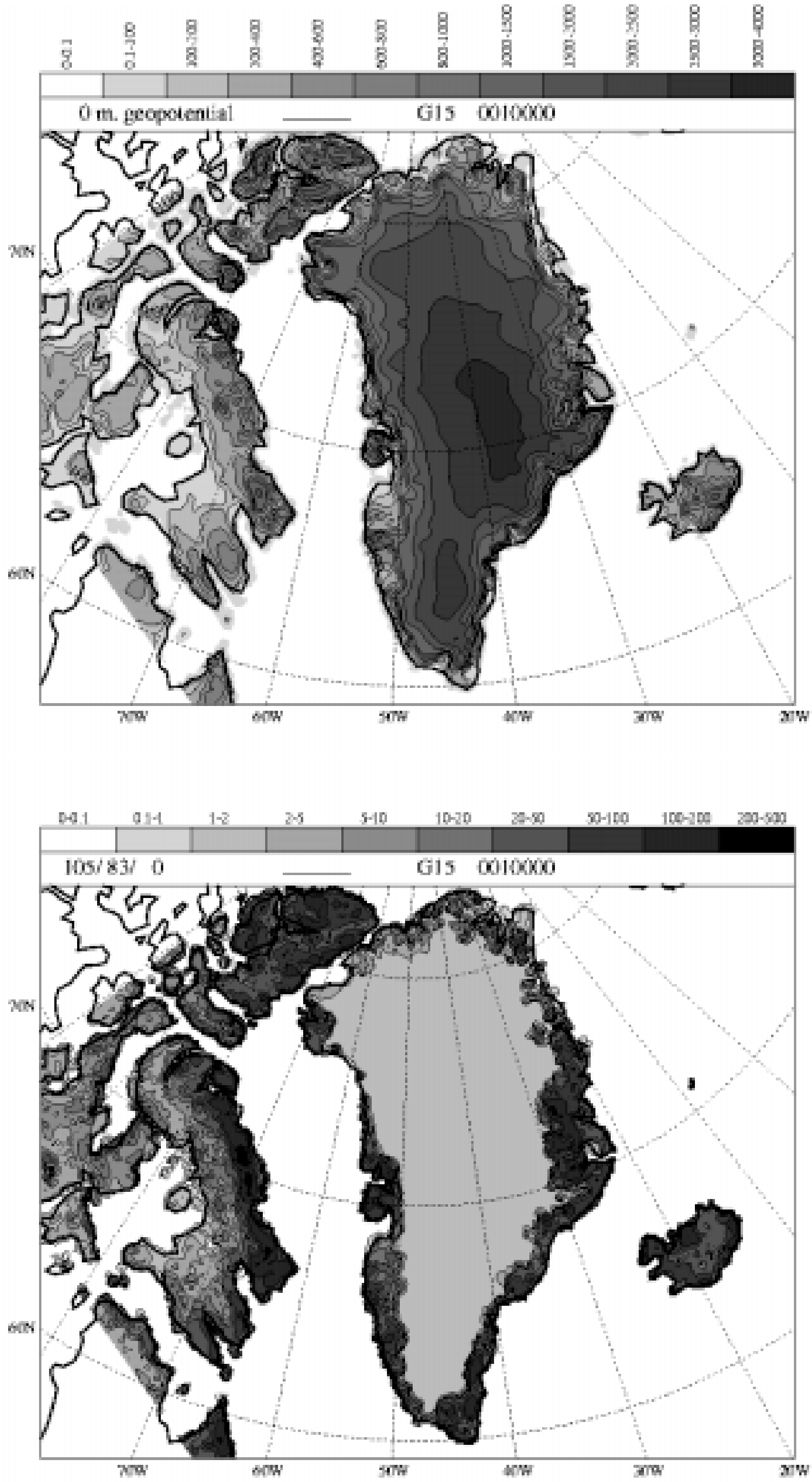


Figure 12: Geopotential (upper) and roughness length (lower) over Greenland as represented in the domain of HIRLAM-N.

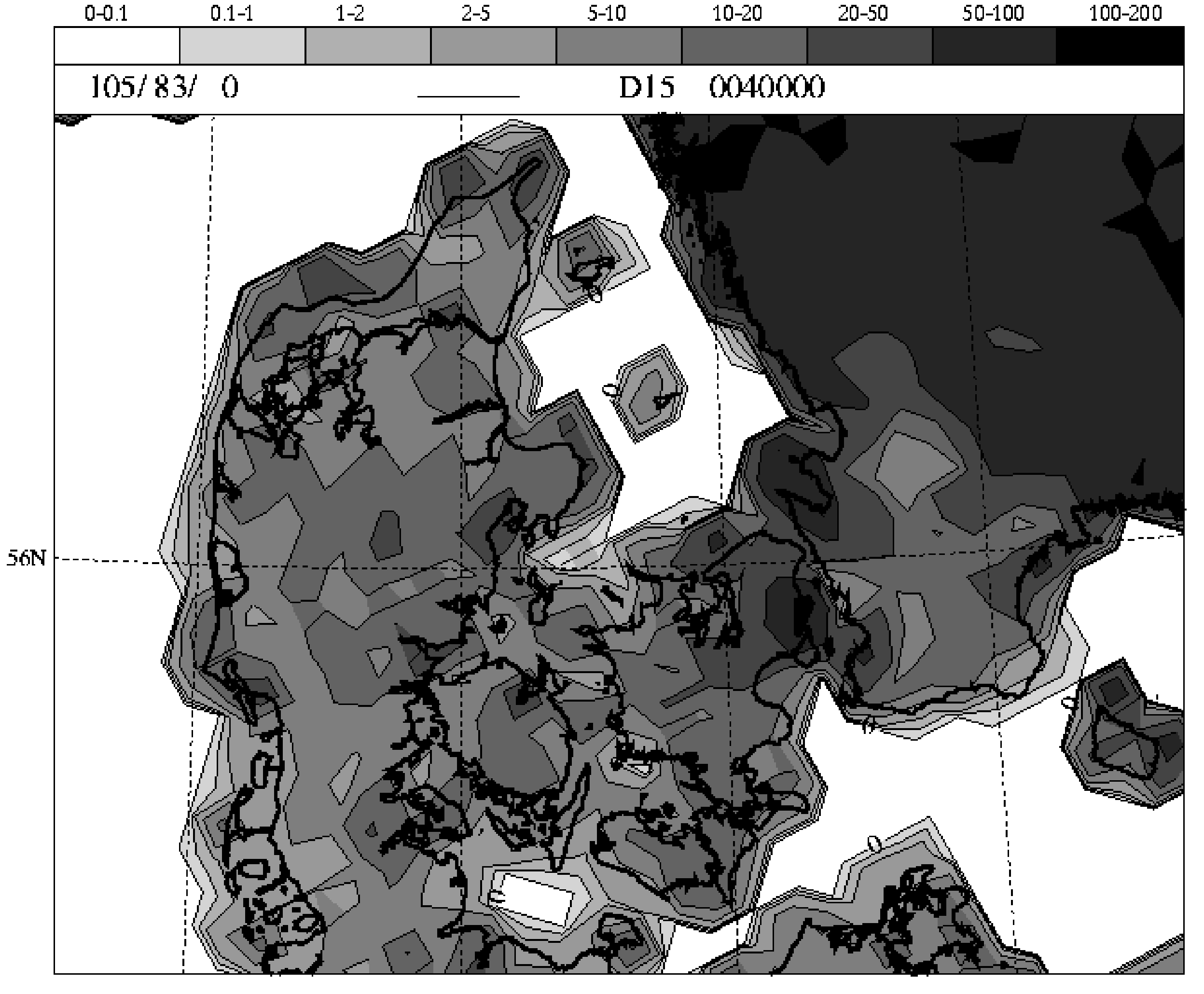


Figure 13: Roughness length for vegetation over Denmark in the representation of the HIRLAM-E domain retrieved with using the landmask. The quality of the coast representation is rather poor.

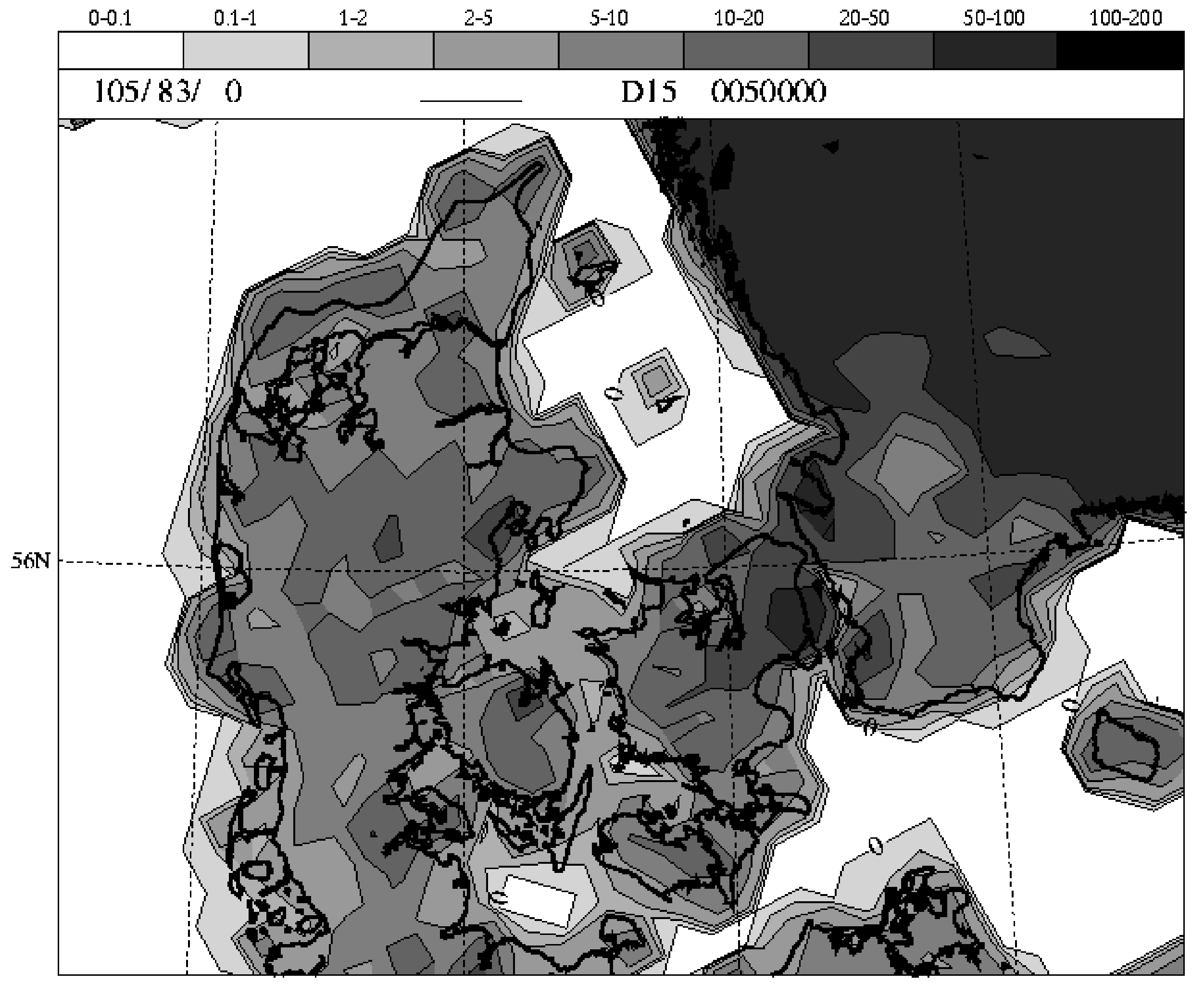


Figure 14: Roughness length for vegetation over Denmark in the representation of the HIRLAM-E domain retrieved with using the landmask and the parameter setting for land use class water. The quality of the coast representation is better than that shown in Figure 13, but it is still poor.

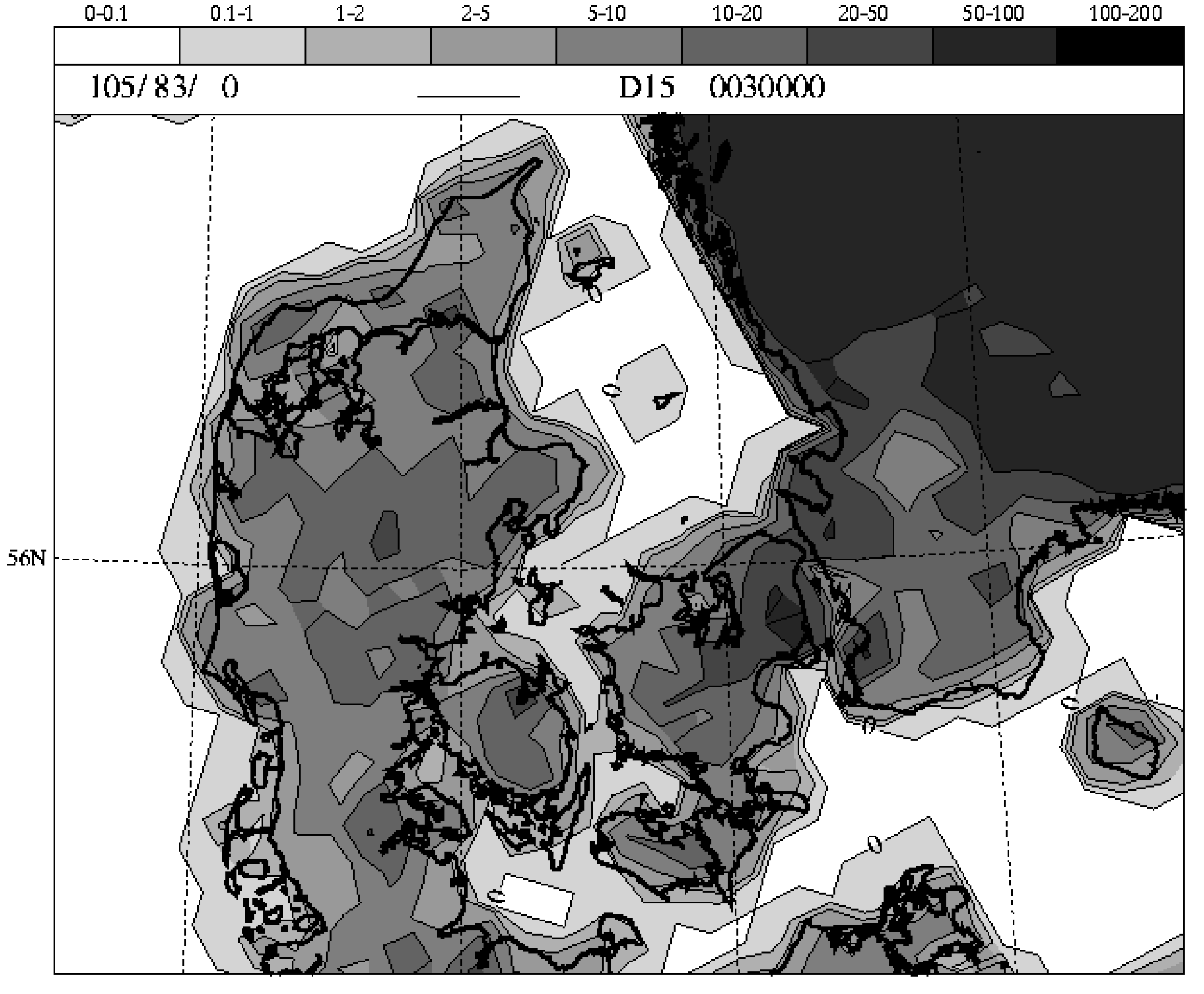


Figure 15: Roughness length for vegetation over Denmark in the representation of the HIRLAM-E domain retrieved only using the parameter setting for land use class water. This gives the best coast representation.

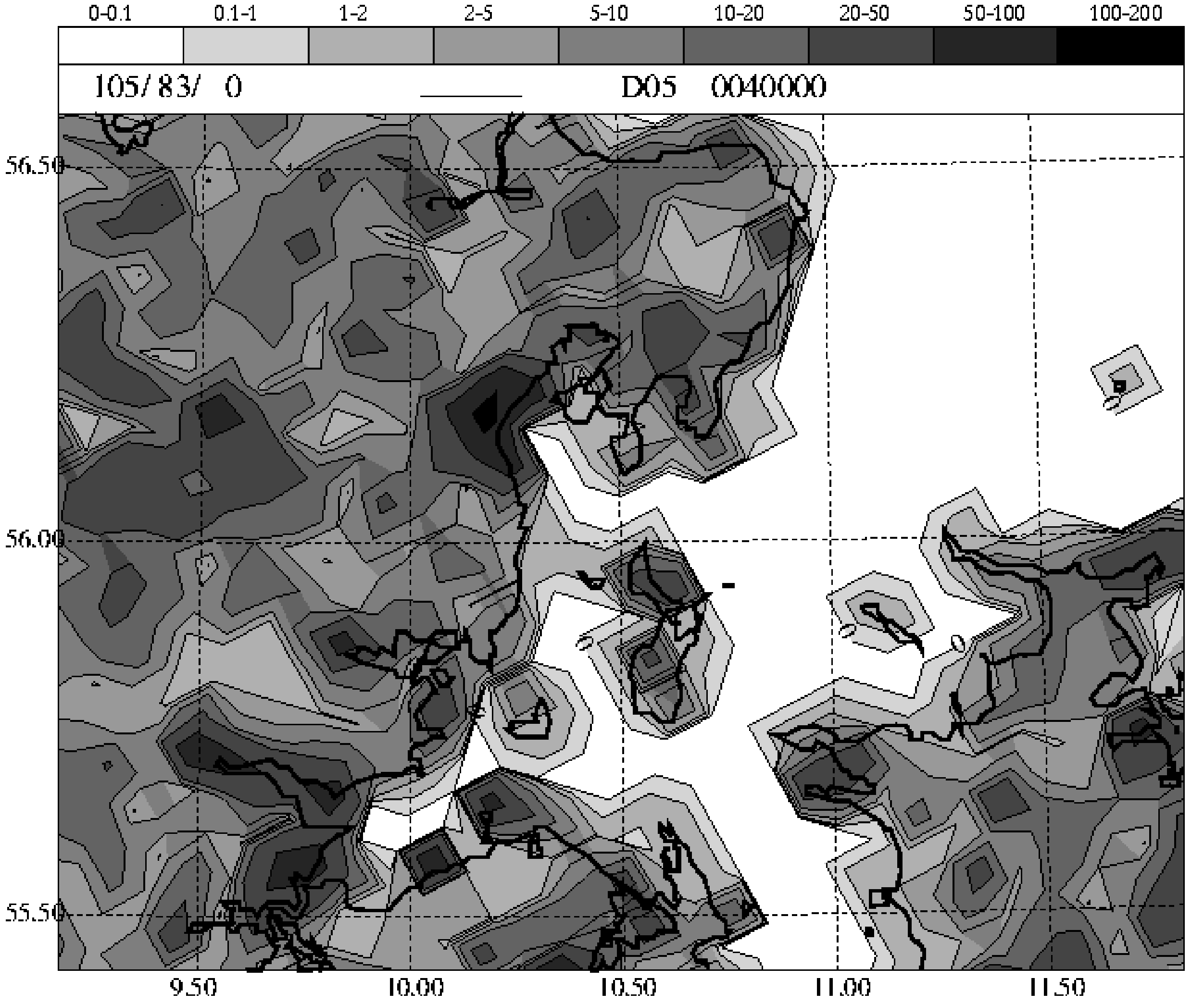


Figure 16: Roughness length for vegetation over Samsø and surroundings in the representation of the HIRLAM-D domain retrieved with using the landmask. The quality of the coast representation is rather poor.

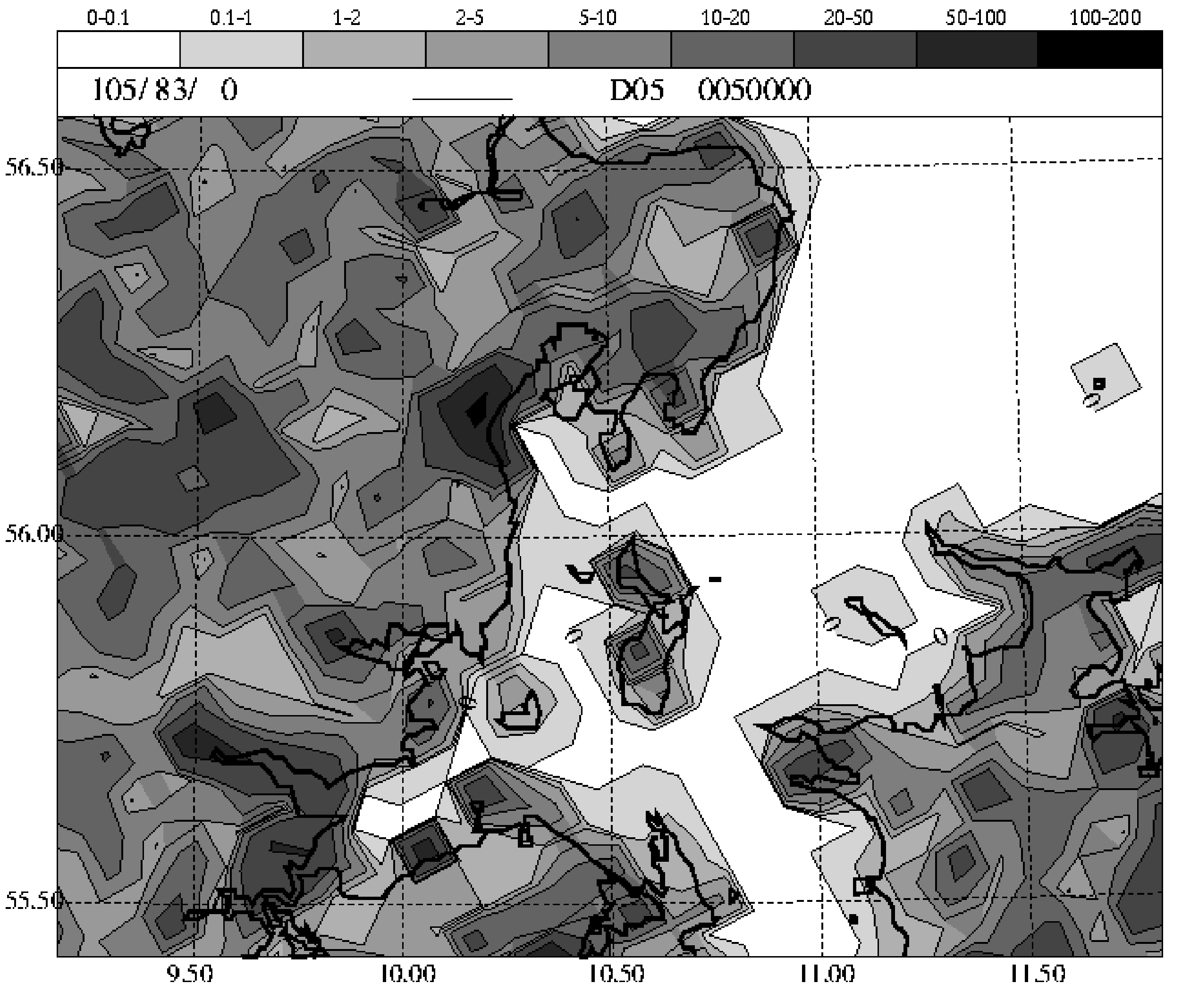


Figure 17: Roughness length for vegetation over Samsø and surroundings in the representation of the HIRLAM-D domain retrieved with using the landmask and the parameter setting for land use class water. The quality of the coast representation is better than that shown in Figure 16, but it is still poor.

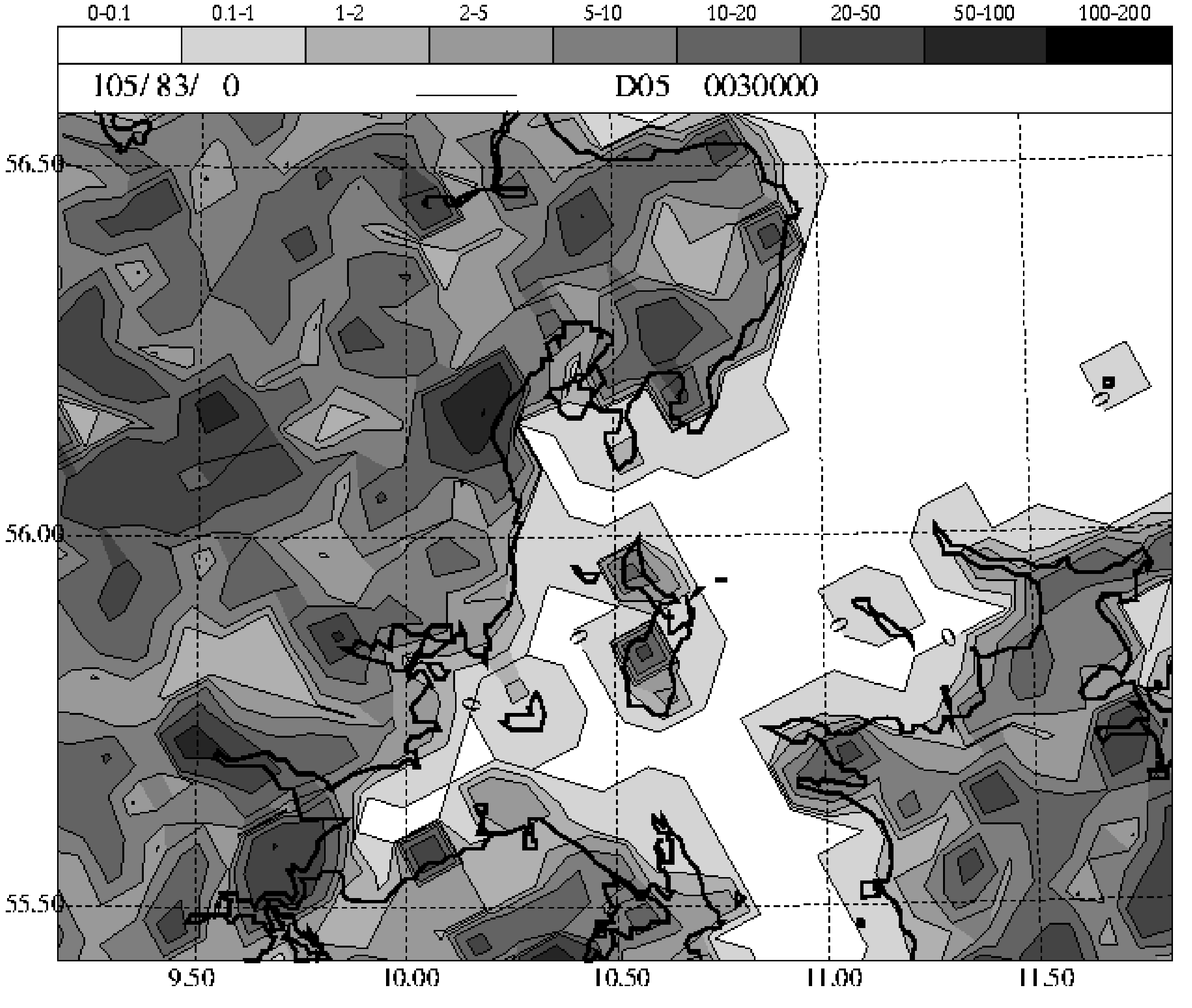


Figure 18: Roughness length for vegetation over Samsø and surroundings in the representation of the HIRLAM-D domain retrieved only using the parameter setting for land use class water. This gives the best coast representation.