

Final Draft
of the original manuscript:

Schroeder, M.; Lipzig, N.van; Ament, F.; Chaboureau, J.P.; Crewell, S.; Fischer, J.; Matthias, V.; Meijgaard, E.van; Walther, A.; Willen, U.:

Model predicted low-level cloud parameters: Part II: Comparison with satellite remote sensing observations during the BALTEX Bridge Campaigns

In: Atmospheric Research (2006) Elsevier

DOI: [10.1016/j.atmosres.2005.12.005](https://doi.org/10.1016/j.atmosres.2005.12.005)

The representation of low-level clouds in atmospheric models. Part II: Spatial distributions from satellite remote sensing during the BALTEX Bridge Campaigns

Marc Schröder ^{a,*}, Nicole van Lipzig ^b, Felix Ament ^c, Jean-Pierre Chaboureau ^d,
Susanne Crewell ^b, Jürgen Fischer ^a, Volker Matthias ^e, Erik van Meijgaard ^f, Andi
Walther ^a, Ulrika Willén ^g

^a *Institut für Weltraumwissenschaften, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10,
12165 Berlin, Germany*

^b *Meteorologisches Institut, Universität München, München, Germany*

^c *Meteorologisches Institut, Universität Bonn, Bonn, Germany*

^d *Laboratoire d'Aerologie, OMP, Toulouse, France*

^e *Institut für Küstenforschung, GKSS, Geesthacht, Germany*

^f *Royal Netherlands Meteorological Institute, KNMI, DeBilt, The Netherlands*

^g *Rosby Center, SMHI, Norrköping, Sweden*

Abstract

A pressing task in numerical weather prediction and climate modelling is the evaluation of modelled cloud fields. Recent progress in spatial and temporal resolution of satellite remote sensing increases the potential for such evaluation efforts. This paper presents a new methodology to compare satellite remote sensing observations of clouds and output of atmospheric models. We discuss the first applications of this method, namely to cloud cover. The

* Corresponding author. Tel.: +49-30-83852751 / fax: +49-30-83856664.

E-mail address: marc.schroeder@wew.fu-berlin.de (M. Schröder)

comparison is carried out for two scenes from the BALTEX Bridge Campaigns. Both scenes are characterised by low-level, shallow clouds with a substantial amount of liquid water. The cloud cover of five different models, LM, Méso-NH, MM5 (non-hydrostatic models), RACMO2, and RCA (regional climate models) as well as corresponding retrievals from remote sensing observations with MODIS onboard TERRA form the basis of a statistical analysis to compare the data sets. Two different measures are defined for comparison: 1) average properties and 2) single cloud features, with the following objectives: A set of parameters which is suitable for an automated, unsupervised analysis and continuous and fast processing of cloud cover received during long-term studies is identified, and the applicability of these parameters is evaluated. It is shown that the newly developed methodology is useful for evaluation purposes and that the extension of average characteristics with single cloud features helps to avoid ambiguities. In particular, the newly introduced patchiness parameters are able to separate differences between the two scenes on the one hand and between the models and the satellite on the other hand. Méso-NH turned out to have a more patchy structure than the other models and MODIS with many small fragmented clouds. The comparison shows that our method can clearly separate differences in cloud cover, which is useful for the derivation of automated, unsupervised algorithms for evaluation of longer timeseries of model output.

Keywords: Patchiness, Contingency tables, Non-hydrostatic models, Climate models, Satellite remote sensing

1. Introduction

Numerical weather prediction and climate models are essential tools for understanding the hydrological cycle, cloud dynamic processes, large scale cloud radiative processes and the energy

budget of the earth. Hence, the development and examination of evaluation strategies is of great importance in order to assess the reliability of the models and to identify limitations.

In the past, clouds and precipitation were not the main focus in model evaluation. With the development of new measuring techniques available for studies on clouds and precipitation, this is changing (e.g. Yu et al., 1996; Barros and Bindlish, 1999; Klein and Jacob, 1999; Hollars et al., 2004; Hennemuth et al., 2003). Most frequently, averages, correlations and histograms of parameters of interest as well as their differences are utilised for evaluation purposes (e.g. Hollars et al., 2004). Yu et al. (1996) concentrated on the sub-grid model variability to identify best cloud overlap within a model column. Klein and Jacob (1999) focused on the positioning of midlatitude baroclinic systems while Barros and Bindlish (1999) approached evaluation with the help of texture analysis. Besides evaluation efforts, the evaluation techniques can partly be utilised to classify cloud systems. E.g., the texture parameters introduced by Barros and Bindlish (1999) are providing a measure for changes in climate monitoring and forecasting activities. Vila and Machado (2004) introduced a series of parameters capable of characterising the shape and internal structure of convective systems. Further examples for texture and shape analysis in remote sensing are presented by Walther and Bennartz (2004). They used this technique to discriminate frontal from non-frontal contributions in precipitation fields. Ryan et al. (2000) present an evaluation of models with different resolutions (ranging from 5 to 300 km), e.g. concluding that the accuracy of regional climate models depends on large scale forcing.

This study extends the results presented in Van Lipzig et al. (2005). They discussed the temporal evolution of low-level clouds, whereas in this paper, we will discuss the spatial characteristics of the same regional atmospheric models using observations of MODIS onboard TERRA. The five models are LM, Méso-NH, MM5 (non-hydrostatic models), RACMO2, and RCA (regional climate models). The major goal is to outline a way to compare cloud cover, being a first order

parameter for the earth's energy budget. The proposed methodology is tested using two cases from the BALTEX Bridge Campaigns in order to identify its usefulness for the evaluation of the models with MODIS data. The evaluation of two cases will not reveal any systematic deficiencies. However, the applicability of the methodology can be assessed to prepare a long-term evaluation. The next section introduces the BALTEX Bridge Campaigns, MODIS, and the models. This is followed by a presentation of the methodology which separates between average characteristics and single cloud features. The results are presented in section 4 and discussed in more detail in section 5. Finally, conclusions are provided.

2. Data sources

2.1 The BALTEX Bridge Campaigns

The BALTEX Bridge Campaigns took place around the Cabauw Experimental Site for Atmospheric Research (CESAR), the Netherlands, in August and September 2001 (BBC) and in May 2003 (BBC2). BBC was a joint venture between the European CLIWA-Net project (Crewell et al., 2002) and the German 4D-Clouds project. BBC2 was a cooperation of several institutions across Europe with major contributions from the 4D-Clouds project and the Dutch weather service, KNMI. During both experiments coordinated observations of clouds by various ground based and airborne instruments were carried out. An overview of BBC including introductions to utilised instruments and first results is given by Crewell et al. (2004). In this study we use observations taken on 23 September 2001 and 21 May 2003. Details on the synoptic situation of these days are provided in part I (Van Lipzig et al., 2005).

2.2 Remote sensing observations

Remote sensing observations utilised in this study were carried out with the Moderate Resolution Imaging Spectrometer (MODIS) onboard the U.S.-American TERRA satellite. MODIS takes 3D measurements of radiances: two spatial dimensions (along track through the propagation of the satellite and across track with 1354 spatial pixels, which corresponds to a swath of ~2330 km) and one spectral dimension (36 channels, spanning the visible, near infrared and thermal range). The spatial resolution of MODIS at nadir varies between 0.25 and 1 km, depending on the chosen channel. Several operational NASA algorithms for the retrieval of various atmospheric and cloud products exist. Among them is the MOD05 product (King et al., 1997) which is used in this study. It is accessible via internet (Distributed Active Archive Center, DAAC:

<http://daac.gsfc.nasa.gov>, funded by NASA). MOD05 data includes a cloud mask that has been developed by Ackerman et al. (1997). The cloud mask has a spatial resolution of 1 km.

The cloud mask utilises calibrated and navigated radiance observations in 17 MODIS channels plus ancillary data inputs, like e.g. viewing geometry and a digital elevation model. Clouds are generally characterised by higher reflectances and lower brightness temperature differences than the underlying earth surface. Therefore, the algorithm relies on a set of radiance differences and ratios as well as absolute values. Major uncertainties in discriminating cloudy from clear sky may arise from thin cirrus, low stratus at night and small cumulus due to insufficient contrast. Cloud edges may also be problematic because the pixel may not be completely covered by clouds (Ackerman et al., 1997).

TERRA is a polar-orbiting satellite that achieves global coverage every 2-3 days. It flies over central Europe 1-2 times a day with overpass times around 10:30 UTC. The comparison of MODIS and model cloud cover is carried out for two scenes which had been taken at 10:45 UTC on 23 September 2001 (S1) and at 10:05 UTC on 21 May 2003 (S2).

2.3 Regional atmospheric models

Output of five models is utilised to illustrate the method: the Lokal Model (LM) from Deutscher Wetterdienst, the MESOScale Non-Hydrostatic model Méso-NH developed by Centre National de la Recherche Scientifique and Météo France, the Mesoscale Model MM5 from the Penn-State University and the National Center of Atmospheric Research and operated by GKSS, the Regional Atmospheric Climate Model RACMO2 developed by KNMI, and the Rossby Center Atmospheric Model (RCA). LM, Méso-NH, and MM5 are non-hydrostatic models while RACMO2 and RCA are hydrostatic limited-area models used for the purpose of regional climate prediction. The models are introduced in detail in Van Lipzig et al. (2005). In this paper, also details about the cloud schemes that are used can be found. In Méso-NH, LM, and MM5, an all-or-nothing scheme is used assuming that clouds are resolved by the model grid. Additionally, LM utilises a subgrid cloud scheme which contributes only to radiation and not to the hydrological cycle. In the climate models it is assumed that the models do not resolve the clouds, and a partially cloud cover is used. The domain size of each model (together with resolution) is given in Fig. 1. The models were initiated at 12 UTC and integrated over a period of 36 hours. 2D spatial fields are recorded every full hour, beginning 12 hours after initialisation. Consequently, each model provides 2D cloud cover at 11:00 (RCA at 10:45 UTC) and 10:00 UTC for S1 and S2, respectively. Although the time difference of 15 minutes can cause significant discrepancies for individual clouds, we do not expect that the statistics of the cloud fields have changed.

3. Approach

The comparison of LM, Méso-NH, MM5, RACMO2, and RCA to MODIS is carried out for cloud cover observations at two specific times without averaging in time. The outermost eight pixels of each model are not considered for the comparison.

The first step of the comparison includes the interpolation of the satellite data to each model grid. If the model utilises a rotated grid, the geographic parameters of MODIS are rotated using the same parameters for rotation as the model does. The transformation follows the equations provided by Pearson, 1990 and is applicable for the LM, RACMO2, and RCA. In the remaining cases, the geographic information of MODIS is interpolated to the model grid by triangulation. If MODIS does not cover the model domain size completely (RACMO2 and RCA), only overlapping areas are considered for comparison. All models have a lower resolution than MODIS. Therefore, the cloud cover is aggregated by straightforward spatial averaging, and a fractional cloud cover is introduced. Histograms of fractional cloud cover are investigated in section 4.1.

In general, fractional cloud cover is kept as long as the determination of average characteristics and single cloud features (see next sections) allows it. As soon as an interpretation of the pixel being either cloudy or cloud free is required, a threshold of 0.5 is defined for the models (LM, RACMO2, RCA) and MODIS: All pixels with fractional cloud cover larger than 0.5 are considered as cloudy and the remaining pixels as cloud free. The chosen threshold is reasonable in the sense that it divides the range of values in equal parts. A histogram analysis of fractional cloud cover does not exhibit specific features, i.e. a clear minimum in the frequency of occurrence, in neither model nor MODIS.

We first investigate the spatial average characteristics (section 3.1) and then focus on single cloud characteristics (section 3.2): The first part provides average information of the cloud cover, e.g. total cloud cover, while the second part describes the features of single clouds in assigning a

certain set of parameters, like e.g. fragmentation, to each cloud. For this analysis, we need to define a cloud entity. Here, a single cloud area is determined using an eight-connected neighbour algorithm: The algorithm searches for cloudy pixels in x- and y- as well as in diagonal direction of the data field to identify connected pixels (see Fig. 2). Each connected cloud area is labelled with a unique area index. If cloud free areas are considered, a four-connected neighbour algorithm is applied in order to avoid pixels being connected across the links of cloudy pixels. In the following, the cloud mask is retrieved from (fractional) cloud cover through application of the threshold, and a single connected cloud area is named ‘cloud’.

3.1 Spatial average characteristics

In this section (spatial) average properties of the cloud cover are presented.

The first to mention is the total cloud cover, b , which is determined by normalising the summation of fractional cloud cover to the total number of pixels.

Second, the quality of the spatial matching between the MODIS and the model cloud masks is assessed. Both cloud masks are simply added and subtracted to construct a contingency table.

This is a 2x2 matrix that, in this work, contains percentages of the number of pixels that are: 1) cloudy in the model and the satellite data, [1,1], 2) cloud free in both cases, [0,0], 3) cloud free in the satellite and cloudy in the model, [0,1], and 4) as 3), but vice versa, [1,0]. The normalisation is carried out to the total number of cloud free and cloudy MODIS pixels, respectively. Overlap plots are generated by assigning different colours to the categories of the contingency table and allow a quick visual impression of possible discrepancies between MODIS and model in terms of overall performance and location of cloud and cloud free areas. The fraction of the pixels that are either zero or one in both fields, i.e. the sum of the pixels corresponding to the [1,1]- and [0,0]-category and normalised to the total number of pixels, defines the *overlap* parameter and can be

understood as single parameter summarising the contingency table. The *overlap* parameter extends the evaluation potential of total cloud cover since it can be interpreted as quality measure for spatial matching of cloudy and cloud free pixels. The visual impression of spatial matching in the overlap plots is expressed in a single parameter. If, e.g. $b=0.5$, the *overlap* can be zero, if the cloud is located in the western part of the MODIS and in the eastern part of the model image. A large number of parameters can be extracted from a contingency table, as discussed, e.g. in Conner and Petty (1998) and Ebert (2003), but our definition of the *overlap* parameter differs from usual ones: The definition of correctly identified pixels includes the [1,1]- and [0,0]-categories, instead of [1,1]-category alone (usually denoted as “hit rate”), because a correct interpretation of cloudy pixels is not of higher priority than one of cloud free pixels.

In the following, the term patchiness is introduced. Generally speaking, the degree of patchiness increases with increasing number of cloud and cloud free areas (N_{cld} and N_{free} , respectively). In Fig. 3, a visual impression is given: The cloud mask in the left panel has a larger degree of patchiness than the cloud in the right panel due to the cloud free pixels within the large cloud, and the necessity to include cloudy and cloud free areas becomes obvious. In order to determine the degree of patchiness, N_{cld} and N_{free} are determined for the MODIS and model cloud masks. The patchiness is then characterised by two parameters: The first parameter gives the overall patchiness of the cloud mask and is defined as

$$p1 = (N_{cld} + N_{free}) / n \quad (1)$$

with n being the total number of pixels. The normalisation is carried out in order to allow a comparison of $p1$ between the different models and MODIS. The degree of patchiness is large if $p1$ is large, and maximum patchiness is achieved, if the cloud mask appears as a chess board with each pixel surrounded by four cloud free and four cloudy pixels. If this is the case, then all overcast pixels are connected and count as a single cloud, which leads to the maximum

$p1 = 1/2 + 1/n$. The second parameter assigns a qualitative attribute to $p1$, and is determined as the normalised difference between N_{cld} and N_{free} :

$$p2 = (N_{cld} - N_{free}) / n. \quad (2)$$

$p2$ can take values between -0.5 and 0.5 . The margins can be explained with the same argument used for the margins of $p1$. If $p2$ is close to -0.5 , then many cloud free areas and only a few clouds are present. The other extreme, a value close to 0.5 , is given, if many clouds, but only a few cloud free areas are found. A value of zero refers to the case, that the number of clouds is identical to the number of cloud free areas. Therefore, $p2$ shows if either cloudy or cloud free areas contribute most to the patchy appearance of the cloud mask. Note, that both values do not change if identical areas are added, i.e. the domain size is duplicated. If $p1$ and $p2$ are determined for the cloud masks in Fig. 3, the largest patchiness is found in the left panel of Fig. 3 ($p1 = 11/150$), dominated by cloud free areas ($p2 = -1/150$), while the cloud mask in the right panel is dominated by clouds ($p1 = 6/150$ and $p2 = 4/150$).

A parameter with similar character as the patchiness, is the contrast which is defined here as the normalised summation of the perimeter (see next section) of all clouds. The normalisation is carried out to the total number of pixels. In contrast to patchiness its dependence on small clouds is less pronounced. The contrast is a measure for the number of boundaries within an image (Barros and Bindlish, 1999).

For identification of these spatial average characteristics, single cloud features need to be defined as is done in the following section.

3.2 Single cloud features

In this section the spatial characteristics are extended to features of single connected cloud areas. The parameters presented in the following describe the shape of a cloud (shape descriptors, e.g.

area and perimeter). The original shape may not entirely be reconstructable from these parameters but they should contain enough characteristics of a shape to discriminate it from other shapes. Topological descriptors, like e.g. fragmentation, describe geometrical structures in a qualitative way. They are measures of the similarity of shapes and are mainly oriented in a way that agrees with human intuition. Recall, that connected cloud areas are labelled with a unique index, utilising an eight-connected algorithm.

The area of a cloud, A , is simply the number of pixels of a single cloud. The center of the cloud is understood as the center of mass with all pixels contributing equally, i.e. the cloud is homogeneous. The center of the cloud is determined by taking the arithmetic mean of the coordinates of all its pixels, separately for elements in x- and y-direction.

The perimeter of a cloud is defined as the number of pixels of the cloud's border, i.e. all cloudy pixels that have at least one neighbouring pixel in x- or y-direction that is cloud free.

Fragmentation, f , is determined by counting all cloud free pixels that are in either x- or y-direction surrounded by cloudy pixels of one cloud. The criterion is not strictly in the sense that the cloud free pixel with indices i and j must have a cloudy neighbour at $i \pm 1$ or at $j \pm 1$ but at any point in the x- or y-direction (see cloud on the right in Fig. 2), as long as it belongs to the same cloud. f is normalised on A and includes the number of cloud free pixels within the cloud, h . h is also retrieved and normalised on A . In Fig. 2, a cloud mask is presented to give an example for the determination of f and h : The cloud on the left shows higher fragmentation ($f = 4/5$ and $h = 0/5$) than the cloud on the right ($f = 3/7$ and $h = 1/7$).

If h is subtracted from f , f characterises the fractional degree of the border of the cloud, the “wigglyness”, and h the brokenness within the cloud. If both parameters are considered, they can be understood as a measure of compactness, i.e. how closely packed is the cloud. Even if identical total cloud cover is observed, qualitative differences can be found due to spatial

displacement (recall *overlap* parameter) and compactness. The idea of estimating the fragmentation of a cloud is taken from Vila and Machado (2004) and is extended by introducing h to separate cloud border from interior effects. Vila and Machado (2004) identified the fragmentation (in combination with area) as a useful parameter to estimate the phase of a cloud field.

4. Results

The comparison of the cloud cover of the models and MODIS, interpolated and adapted to each of the model's resolution, is presented in this section. The section is divided into three subsections: In the first subsection cloud cover images are presented, in the second subsection average characteristics of model and MODIS cloud cover and in the third subsection single cloud features of cloud masks are compared.

4.1 Cloud cover

In Fig. 4 the cloud cover observed by MODIS for S1, left panel, and S2, right panel, is shown. The S1 cloud mask is dominated by a large, unbroken cloud over the Netherlands and two relatively large cloud free areas around the coast of the Netherlands and at the SE corner of the image. In contrast, the S2 cloud mask is characterised by many relatively small cloud free areas, in particular over land, which reflects the convective nature of this day.

The cloud cover of the five models is presented in Fig. 5 (S1) and Fig. 6 (S2). The same area is shown in all panels (reduced domain size of the climate models). All models are characterised by a relatively large amount of total cloud cover with MM5 having smallest total cloud cover. The MM5 cloud cover found for S2 shows a clear-sky North Sea and large cloud free areas over western Netherlands. Also LM and RACMO2 show a part of the North Sea cloud free, but the

agreement with MODIS is better than for MM5. The convective nature of S2 becomes evident in the cloud cover of LM and Méso-NH: They show many relatively small cloud free areas.

The agreement between cloud cover of the climate models and MODIS for the full domain is excellent. The major difference is a larger cloud cover over the Baltic Sea for RCA in comparison to MODIS and RACMO2 (not shown).

The number of pixels with fractional cloud cover different from zero and one is between 10% and 20%, if non-hydrostatic models are considered. In contrast, the amount of such pixels is significantly larger (up to 50%) in case of the climate models. Obviously, the probability of a pixel being either cloud free or cloudy decrease with increasing pixel size (confirmed by Tian and Curry, 1989 and Willen et al., 2005), simply because small-scale cloud structure is no longer resolved. If histograms of fractional cloud cover (not shown) of MODIS are compared to the models (LM, RACMO2, and RCA), they exhibit the following: The three models underestimate the frequency of occurrence of cloud free pixels, and the climate models are more skewed than than the satellite observations. This was also found by Willen et al. (2005).

4.2 Comparison of spatial average characteristics of MODIS and model cloud cover

Average characteristics and contingency tables are determined for an area that is covered by all models (roughly between 3°E and 7.5°E as well as 51°N and 53.5°N). However, the overlap plots show the full domains of LM, Méso-NH, and MM5 (and reduced domains of RACMO2 and RCA).

Tables 1 and 2 give the total cloud cover, the *overlap*, and the patchiness for the models and MODIS for S1 and S2, respectively. The total cloud cover is well represented by most models, within about 10% of the observations, and the relatively large *overlap* shows that the spatial distributions of clouds and cloud free areas are comparable. The largest difference to MODIS is

found for MM5, which underestimates the S2 total cloud cover by ~25%. This is related to the large cloud free region over the North Sea and western Netherlands which is not present in MODIS observations (see previous section). Also RACMO2 underestimates the S2 total cloud cover, but the location of the cloud free area over sea, just west and north of the Netherlands, is well represented. This is reflected in the *overlap* parameter, which is about 13% larger for RACMO2 than for MM5.

Analysing patchiness revealed the following: The patchiness in LM corresponds most closely to the measurements for these two cases, although the patchiness is underestimated in both scenes. In contrast to LM, Méso-NH overestimates the patchiness in both scenes: The number of clouds and cloud free areas is much larger than the measurements indicate. As explained earlier, the difference in number of clouds and number of cloud free areas is relevant for $p2$. From $p2$ it is clear that there are too much cloud free areas in this model, rather than too much clouds. The overestimation of the occurrence of small cloud free areas can also be seen in the histogram analysis of area in section 4.3. The overestimation of patchiness in case of Méso-NH might be related to the fact that shallow convection is not parameterised in this model. In addition, no subgrid scale cloud scheme is used in Méso-NH. On the other hand, LM does not represent shallow convection either, and the patchiness in this model is underestimated for the two cases. The MM5 results are similar to those of LM for S1, but its patchiness for S2 shows largest observed underestimation: There is one large cloud situated over most of the Netherlands, whereas the North Sea and the west of the Netherlands are cloud free. In part I of this paper, we have already seen that for S2, MM5 overestimates the lifetime of clouds over Cabauw, which is reflected in an underestimation of the variability in liquid water path, and probably explains the low patchiness.

In both scenes the patchiness related to the climate models is larger than the patchiness related to the non-hydrostatic models. This is also found for MODIS after interpolation to either the climate or the non-hydrostatic model grids. This is a consequence of the very different total number of pixels (given in the caption of Table 1) between both groups of models: Small changes in N_{cld} or N_{free} would dramatically change the patchiness if n is small. One might argue that as long as the total number of pixels is large for both groups (but nevertheless very different) a comparison between models of different resolutions is possible, i.e. the maximum $p1$ is close to 0.5 for both cloud masks (see discussion for Eq. (1)). But this is only given if the sum of N_{cld} and N_{free} increases in the same way as the total number of pixels does. This will strongly depend on the considered cloud cover situation and cannot be expected a priori. However, for evaluation purposes, i.e. comparison between satellite and model observations, patchiness is well suited. In Fig. 7 the overlap plots of the five models are presented for S1 and allow a direct visual impression of the performance of the models. In general, the overlap plots show large spatial agreements, mainly due to a large cloud found in both MODIS and model cloud masks. All models consistently have larger cloud cover than MODIS at the west coast of the Netherlands. Table 3 quantifies the results of the overlap plots (recall also the *overlap* parameter in Table 1). Of all pixels, the pixels that are cloudy in the model and in MODIS ([1,1]) have the largest number of occurrence. In addition, there are more pixels where MODIS is cloud free and the models have a cloud ([0,1]) than vice versa ([1,0]). A good performance in view of the contingency table is found for LM and RACMO2. MM5 shows smallest percentage in the [1,1]-category and largest in the [0,0]-category. The latter is a consequence of an underestimation of total cloud cover by MM5 what increases the probability for coincidence with MODIS cloud free pixels. If the percentages are achieved by normalisation on total number of pixels, we find 52% in the [1,1]- and 11% in the [0,0]-category for MM5. However, the latter value is not

representative since it depends on $1 - b$. The *overlap* parameter is determined by normalisation on total number of pixels and provides information on the spatial matching of cloudy and cloud free pixels (recall section 3.1). The contingency tables is an extension of this measure and shows details on contributions from cloudy and cloud free pixels.

Overlap plots for S2 are provided in Fig. 8. LM and Méso-NH, each by a different degree, reproduce the patchy appearance of the spatial distribution of clouds over land surfaces. MM5 shows a few cloud free patches, but less than MODIS (and LM and Méso-NH). If the climate models are considered, RCA shows patchy structure over the Netherlands but MODIS and RACMO2 not. Table 4 summarises the performance of the models. In all cases besides MM5, the largest percentages are found for the [1,1]-category and the largest misinterpretations for the [0,1]-category. Remarkable is the high percentage for RCA in the latter case. In Fig. 8e it can be seen that cloud free areas are shifted in NE direction between RCA and MODIS.

Average parameters, overlap plots, and contingency tables (not shown), retrieved from the full domain size of RACMO2 and RCA, are also analysed. Both models underestimate total cloud cover but have a large *overlap*. The patchiness is smaller, if compared to MODIS and the reduced domain results, and still dominated by cloud free areas. In contrast to the reduced domain, the contingency tables show significantly larger values in the [0,0]-category. If the comparison of histograms of cloud cover is recalled (section 4.1), the climate models underestimate clear-sky. This is compensated by larger frequencies at values between 0 and 0.5.

4.3 Comparison of single cloud features of MODIS and model cloud masks

The cloud cover of satellite and models are dominated by a few large clouds, recall section 4.1 and results related to p_2 , and all other clouds are significantly smaller in size. Therefore, single

cloud features are determined for cloud free areas. Furthermore, they are determined for the full domain of each model to optimise the variability of classes and frequency of occurrence.

In Fig. 9 area and in Fig. 10 fragmentation histograms of the models and MODIS are given for S1 and in Fig. 11 and Fig. 12 for S2. All models underestimate the frequency of occurrence of small cloud free areas ($A < 10$) and the frequency of small fragmentations ($f < 0.3$), except Méso-NH: The cloud mask of Méso-NH consists of more small areas and spans a wider variety of fragmentation than MODIS does.

Although Méso-NH largely overestimates the patchiness and differences to histograms related to MODIS are quite large, it has the most realistic representation of the frequency of occurrence of cloud free area and fragmentation among the non-hydrostatic models. Many small cloud free areas can be identified for S1 and S2, and the frequency of occurrence decreases with area. Also small fragmentations have a high frequency of occurrence, and this frequency decreases with increasing f . From this it is clear that one single measure is not sufficient to have a good description of the spatial distribution of clouds, and by looking at different measures, a more complete view can be obtained.

All models except Méso-NH underestimate the frequency of occurrence of small clouds. When a cloud is formed, it is covering a larger area than the measurements indicate, at least for the two scenes considered. From a comparison of measured time series of liquid water path at Cabauw (Van Lipzig et al., 2005), it turned out that the lifetime of clouds is underestimated by LM. An analysis of its model output pointed to a large variability in the vertical velocity occurring together with a large variability in the liquid water path. Apparently, this development and disappearance of clouds occurred in a relatively large area around Cabauw. This is in contrast to Méso-NH, where the lifetime of clouds is underestimated, but the area of the cloud is also underestimated.

The climate models underestimate the frequency of occurrence of small clouds. This is consistent with Van Lipzig et al. (2005), where it is shown that variability in the liquid water path is smaller than the measurements indicate, even after aggregating the measurements to the model grid. The remarkable large fragmentations for the climate models for S1 result from a large v-shaped cloud free area in RACMO2 and RCA, which is not connected in the MODIS cloud mask, while the large fragmentation of MODIS in Figs. 12d and e has its origin in the cloud free area over the Baltic Sea which is less compact than in the models.

The variety of cloud areas and fragmentations as well as their frequency of occurrence should be larger, i.e. a larger number of cases is needed to increase the reliability of conclusions. However, the histograms enabled the identification of discrepancies between model and MODIS and are helpful to aid the conclusions based on patchiness.

Histograms of h are not shown since their frequency and variability is not sufficient. It shall be mentioned here, that for MODIS and in particular for Méso-NH h is larger for S2 than for S1.

Histograms of h could be seen as a complement to patchiness: h is given in pixel numbers while patchiness is defined as number of patches, being either cloud free or cloudy and regardless of their area.

5. Discussion

This section shows differences between the two scenes, and deepens the discussion about the usefulness of the introduced methodology.

The comparison of total cloud cover of MODIS between S1 and S2 (see Tables 1 and 2), shows a difference of about 10% which is not reflected in the models. The reason can be found in the large cloud free area at the west coast of the Netherlands observed by MODIS which is not present in the models.

The average parameters for the satellite are not identical in the first three columns in Tables 1 and 2, even though the considered domain is the same in all three cases. This can in part be explained by different resolutions (see caption of Fig. 1): Fig. 4 reveals that many small cloudy and cloud free areas are found in the satellite image, so that even though the difference in the resolutions between the non-hydrostatic models is small, the presence of small areas seems to have an effect on the difference, at least for the patchiness. A second reason can be interpolation uncertainties due to different resolutions, interpolation approaches, and geographic coordinates. This gives an indication of the robustness of the measures, and it can be concluded that differences in patchiness (in total cloud cover) of the order of 0.2×10^{-2} (0.02) are not significant. This does not affect previously made conclusions.

The contrast parameter can be utilised to confirm the results related to patchiness. It has the advantage that it does not depend as strongly as the patchiness parameters on small cloud and cloud free areas, and therefore, it does not depend so much on the resolution of the models and interpolation uncertainties. The results related to contrast are qualitatively similar to those related to patchiness, but show smaller variations between MODIS data related to the different models. Fig. 4 exhibits a remarkable difference between cloud cover for S1 and S2: Convection results in large spatial variabilities in cloud cover over land surfaces for S2 while the S1 cloud cover is more homogeneous, caused by a large stratiform cloud. If the patchiness parameters are compared between both cases, neither in the MODIS nor in the non-hydrostatic model results this visual impression is reflected in larger patchiness for S2. The reason for this is that many broken clouds both in the models and in MODIS are found at the coast and over the North Sea for S1 (verify Figs. 4, 5, and 6). If the patchiness parameters are calculated over land and for longitudes larger than 5°E , MODIS as well as LM and Méso-NH observe significantly larger absolute values for S2 than for S1 but not MM5 (not shown). This points to the conclusion that LM and

Méso-NH are better in representing the convective cloud structure for S2 than MM5, which is in agreement with Van Lipzig et al. (2005).

The analysis of scatterplots of A versus f and h shows neither clusters nor relevant correlations.

An exception is LM (S1) where a slight dependence of fragmentation on area could be found for large areas. This might explain the smooth appearance of the cloud mask of LM since cloud free areas partly surround a few but large cloud segments. A similar argument was given for the large fragmentations found in the histograms of fragmentation of RACMO2 and RCA (see section 4.3).

Another aspect is that in particular for small clouds, minor changes in the cloud mask can largely affect fragmentation: If the cloudy pixel at the lower right corner of the cloud on the right in Fig. 2 would not be present, $f = 1/7$ (and $h = 0$) is found, a change by a factor of three. Therefore, histograms of fragmentation should be accompanied by histograms of area.

There are several problems related to an evaluation of the cloud cover alone. Such an evaluation does not give any information about cloud optical properties and vertical distribution of cloud layers. Therefore, the method presented here should be extended to cloud optical thickness, being related to liquid water path, to cloud top heights, and cloud phase.

Furthermore, the comparison was carried out for two cases of specific time, and the temporal evolution of clouds was not considered in this paper. However, this paper adds spatial aspects to the temporal evolution of clouds presented in part I (Van Lipzig et al., 2005). The combination of both studies helped to draw conclusions from both parts.

6. Conclusions

In this second part of a series of two papers (first part by Van Lipzig et al., 2005), several methods are presented for a quantitative evaluation of cloud patterns in atmospheric models, namely LM, Méso-NH, MM5, RACMO2, and RCA. The usefulness of the methodology is tested

by applying it to cloud cover observations at 10:45 UTC on 23 September 2001 (S1) and at 10:05 UTC on 21 May 2003 (S2). Patchiness and probability density functions of the area of clouds turned out to be very relevant to distinguish different types of cloud cover. For example, the patchiness enabled the differentiation between the convective cell structures observed for S2 and the more stratiform cloud cover for S1, if restricted to land surfaces. Furthermore, Méso-NH turned out to have a different spatial distribution of clouds for both S1 and S2 compared to the other non-hydrostatic models and MODIS. Already by looking at the cloud cover, this difference between the models can be clearly identified, with Méso-NH showing a lot more small scale structure. This is reflected in an overestimation of the patchiness parameter ($p1$) compared to the MODIS results. The analysis of patchiness and histograms of area and fragmentation revealed, that a single measure is not sufficient to have a good description of the spatial distribution of clouds. The average characteristics need to be complemented by the single cloud features, e.g. the interpretation of patchiness is easier, if area histograms are provided.

Based on two scenes, systematic deficiencies in modelled cloud fields cannot be identified. However, the overestimation of the patchiness of the two scenes in Méso-NH is consistent with a short lifetime of clouds in this model at Cabauw (Van Lipzig et al., 2005). Apparently, small cloud structures are advected over the site. This result is in contrast with LM: Although Van Lipzig et al. (2005) showed that also LM underestimates the lifetime of clouds at Cabauw, correspondence with MODIS is good, with a tendency of LM to underestimate patchiness. So clearly, the short lifetime of clouds at Cabauw is not due to the fact that the cloud structures are too small, but rather related to appearance and disappearance of clouds on short time scales. This is confirmed by comparing the vertical velocities in LM, that are also highly variable in time and correlated with the liquid water content. This example illustrates how the combination of

information on spatial and temporal distribution of clouds can help to gain insight into the behavior of the models.

This paper has provided a method by which satellite data can be used for model evaluation and shows results of first applications to the BBC cases – one of the cases of the WMO cloud modeling workshop 2004. The description of the spatial distribution in the model is complemented by a description of the synoptic situation and temporal evolution of shallow clouds at Cabauw in part I of this paper (Van Lipzig et al., 2005).

For an atmospheric model it is not feasible (nor relevant) to forecast an individual cumulus cloud system at the exact location and time. However, the models need to be able to describe the statistical properties of the cloud cover. Satellite images are of particular use since they cover a large area at high spatial resolution.

In the near future the evaluation of non-hydrostatic and climate models with satellite remote sensing should be extended to a significantly larger amount of cases, i.e. a long-term study is planned. In particular, this study will allow the analysis of the diurnal cycle of cloud systems, if data from the MSG satellite, which takes full disk observations every 15 minutes, is utilised. Furthermore, the comparison should be extended to other cloud properties, like optical thickness, cloud top height, and cloud phase. The methodology presented by Walther and Bennartz (2004) will be used to process satellite data and model output in order to separate frontal from non-frontal cloud systems. If a tracking technique is applied to two successive images, separately, for the satellite and the model, it will offer a way to identify and compare the path of frontal systems. Another relevant study will be the analysis of the mismatch between satellite and model in, e.g., the vicinity of frontal areas, using the adopted frontal/non-frontal classification and the overlap plots. The latter might also be useful to find areas where misinterpretations, either over- or

underestimations, are most frequent and which might be related to, e.g., orography or land/sea transitions.

The parameters and methods presented above allow a fast, unsupervised processing of data sets for evaluation purposes and therefore are applicable to long-term evaluations with significantly larger data amounts.

Acknowledgements

The authors like to thank Peter Albert from Deutscher Wetterdienst and Maximilian Reuter from the Freie Universität Berlin for providing their interpolation routines. The use of MM5, which is a Pennsylvania State University/National Center for Atmospheric Research (NCAR) model, is gratefully acknowledged. N.P.M. van Lipzig and M. Schröder were funded by the German project QUEST under grant FI 435/12-1 and CR 111/5.

References

Ackerman, S., Strabala, K., Menzel, P., Frey, R., Moeller, C., Gumley, L., Baum, B., Schaaf, C., Riggs, G., 1997. Discriminating clear-sky from cloud with MODIS: Algorithm Theoretical Basis Document (MOD35). Algorithm Theoretical Basis Document ATBD-MOD-06, NASA Goddard Space Flight Center, 125 pp.

Barros, A.P., Bindlish, R., 1999. Using image analysis techniques for intercomparison of spatial variables: An application to satellite observations and model simulations of cloud fields. *Systems Research and Info. Systems* 8, 273-293.

Crewell, S., Drusch, M., van Meijgaard, E., van Lammeren, A., 2002. Cloud observations and modelling within the European BALTEX Cloud Liquid Water Network. *Boreal Environment Research* 7, 235-245.

Crewell, S., Bloemink, H., Feijt, A., Garcia, S., Jolivet, D., Krasnov, O.A., van Lammeren, A., Löhnert, U., van Meijgaard, E., Meywerk, J., Quante, M., Pfeilsticker, K., Schmidt, S., Scholl, T., Schröder, M., Simmer, C., Trautmann, T., Venema, V., Wendisch, M., Willen, U., 2004. The BALTEX Bridge campaign: An integrated approach for a better understanding of clouds. *Bull. Amer. Meteor. Soc.* 85, 1565-1584.

Conner, M.D., Petty, G.W., 1998. Validation and intercomparison of SSM/I rain-rate retrieval methods over the continental United States. *J. Appl. Met.* 37, 679-700.

Ebert, E.E., 2003. Verifying satellite precipitation estimates for weather and hydrological applications. In Proceedings of the 1st International Precipitation Working Group Workshop, Madrid, Spain, EUMETSAT EUM P34, ISBN 929110-045-5, 215-224.

Hennemuth, B., Rutgersson, A., Bumke, K., Clemens, M., Omstedt, A., Jacob, D., Smedman, A.-S., 2003. Net precipitation over the Baltic Sea for one year using models and data-based methods. *Tellus* 55A, 352-367.

Hollars, S., Fu, Q., Comstock, J., Ackerman, T., 2004. Comparison of cloud-top height retrievals from ground-based 35 GHz MMCR and GMS-5 satellite observations at ARM TWP Manus site. *Atm. Res.* 72, 169-186.

King, M.D., Tsay, S.C., Platnick, S.E., Wang, M., Liou, K.-N., 1997. Cloud retrieval algorithms for MODIS: Optical thickness, effective particle radius, and thermodynamic phase. Algorithm Theoretical Basis Document ATBD-MOD-05, NASA Goddard Space Flight Center, 83 pp.

Klein, S.A., Jacob, C., 1999. Validation and sensitivities of frontal clouds simulated by the ECMWF Model. *Mon. Wea. Rev.* 127, 2514-2531.

Pearson, F. II, 1990. *Map projections: Theory and applications*. CRC Press, Boca Raton, FL.

Ryan, B.F., Katzfey, J.J., Abbs, D.J., Jakob, C., Lohmann, U., Rockel, B., Rotstajn, L.D., Stewart, R.E., Szeto, K.K., Tselioudis, G., Yau, M.K., 2000. Simulations of a cold front by cloud-resolving, limited-area, and large-scale models, and a model evaluation using in-situ and satellite

observations. *Mon. Wea. Rev.* 128, 3218-3235.

Tian, L., Curry, J.A., 1989. Cloud overlap statistics. *J. Geophys. Res.* 94, 9925-9935.

Van Lipzig, N.P.M., Schröder, M., Crewell, S., Ament, F., Chaboureau, J.-P., Löhnert, U., Matthias, V., van Meijgaard, E., Quante, M., Willen, U., Yen, W., 2005. The representation of low-level clouds in atmospheric models. Part I: Temporal evolution from ground-based remote sensing during the BALTEX Bridge Campaigns. Submitted to *Atm. Res.* (this issue).

Vila, D.A., Machado, L.A.T., 2004. Shape and radiative properties of convective systems observed from infrared satellite images. *Int. J. Rem. Sens.* 25, 4441-4456.

Walther, A., Bennartz, R., 2004. Radar-based precipitation classification in the Baltic area. Submitted to *Tellus A*.

Willen, U., Crewell, S., Baltink, H.K., Sievers, O., 2005. Assessing model predicted vertical cloud structure and cloud overlap with radar and lidar ceilometer observations for the Baltex Bridge Campaigns of CLIWA-NET. Accepted by *Atm. Res.*

Yu, W., Doutriaux, M., Seze, G., Le Treut, H., Desbois, M., 1996. A methodology study of the validation of clouds in GCMs using ISCCP satellite observations. *Clim. Dyn.* 12, 389-401.

Fig. 1. Illustration of the different domain sizes of the non-hydrostatic and climate models. The MM5 domain lies completely within the domain of LM. The resolutions of the models are 2.8 (LM), 2.5 (Mésó-NH), 3.0 (MM5), and 19 km (RACMO2, RCA).

Fig. 2. Gimic showing two clouds (white pixels) imbedded in a cloud free area (black). White arrows indicate neighbours in x- and y-direction and black arrows neighbours in diagonal direction. Connected cloud areas are determined using an eight-connected algorithm while connected cloud free areas are identified by a four-connected algorithm. Therefore, the cloud free pixel in the center of the right cloud does not belong to the large cloud free area. f marks pixels contributing to fragmentation and h pixels that contribute to number of cloud free pixels within a single cloud. The total number of pixels in each panel is $15 \times 10 = 150$.

Fig. 3. Gimic showing two cloud covers (white and black pixels indicating cloudy and cloud free pixels, respectively). The cloud cover on the left contains five clouds and six cloud free areas, the cloud cover on the right five clouds and one cloud free area. Therefore, the degree of patchiness is larger in the left than in the right panel.

Fig. 4. Cloud cover retrieved from MODIS S1 (Fig. a) and S2 (Fig. b) data. Cabauw, 4.93°E and 51.96°N , is marked in both panels.

Fig. 5. Cloud cover of LM (a), Mésó-NH (b), MM5 (c), RACMO2 (d), and RCA (e) for S1. Cabauw, 4.93°E and 51.96°N , is marked in both panels. RACMO2 and RCA show a section of their full domain to allow better comparison between the models.

Fig. 6. As Fig. 5 but for S2.

Fig. 7. Overlap plots between MODIS and LM (a), Méso-NH (b), MM5 (c), RACMO2 (d), and RCA (e) observations for S1. The overlap for RACMO2 and RCA is shown for a reduced domain size. The legends provide cloud cover for MODIS (first value) and model (second value). The white and black cruxes mark the center position of the largest [0,1]- and [1,0]-areas. Cabauw, 4.93°E and 51.96°N, is marked in all panels.

Fig. 8. Overlap plots between MODIS and LM (a), Méso-NH (b), MM5 (c), RACMO2 (d), and RCA (e) observations for S2. The white and black cruxes mark the center position of the largest [0,1]- and [1,0]-areas in Figs. (a), (b), (d) and (e). In Fig. (c) both cruxes are white. Besides this, the figure set up is identical to that of Fig. 7.

Fig. 9. Cloud free area histograms of LM (a), Méso-NH (b), MM5 (c), RACMO2 (d), and RCA (e) in comparison to MODIS (S1). Not included are one-pixel-areas. Note the different range for Méso-NH.

Fig. 10. Fragmentation histograms of LM (a), Méso-NH (b), MM5 (c), RACMO2 (d), and RCA (e) in comparison to MODIS (S1). The maximum fragmentation of 1.2 for RACMO2 is not shown. Fragmentation is determined for cloud free areas. Not included is a fragmentation of zero.

Fig. 11. As Fig. 9 but for S2.

Fig. 12. As Fig. 10 but for S2 and with maximum fragmentation of RACMO2 below one.

Table 1

Comparison of average parameters extracted from the S1 data (b : total cloud cover; $p1$, $p2$: patchiness). $p1$ and $p2$ were multiplied by 100 to provide larger values. The values in brackets refer to MODIS. The total number of pixels is between 11000 and 19000 for the non-hydrostatic and almost 300 for the climate models.

Parameter	LM	Méso-NH	MM5	RACMO2	RCA
b	0.83 (0.74)	0.75 (0.73)	0.65 (0.74)	0.69 (0.74)	0.70 (0.74)
<i>overlap</i>	0.76	0.68	0.63	0.75	0.66
$p1 / 100$	0.28 (0.66)	1.24 (0.62)	0.30 (0.65)	1.43 (1.79)	1.43 (1.79)
$p2 / 100$	-0.22 (-0.34)	-0.92 (-0.27)	-0.01 (-0.27)	-0.00 (-0.36)	-0.72 (-0.36)

Table 2

As Table 1 but for S2.

Parameter	LM	Méso-NH	MM5	RACMO2	RCA
B	0.72 (0.84)	0.79 (0.84)	0.59 (0.84)	0.69 (0.85)	0.72 (0.85)
<i>overlap</i>	0.74	0.72	0.67	0.77	0.76
$p1 / 100$	0.24 (0.59)	1.28 (0.56)	0.06 (0.74)	2.51 (1.43)	2.51 (1.43)
$p2 / 100$	-0.21 (-0.55)	-1.13 (-0.50)	-0.05 (-0.69)	-1.08 (-0.72)	-1.79 (-0.72)

Table 3

Contingency tables for the models (S1). The percentages are retrieved by normalisation on the number of cloud free and cloudy pixels as observed by the satellite. The spatial distribution of the classes (e.g. [1,1], first entry in all contingency tables) can be seen in Fig. 7.

<i>LM / MODIS</i>	cloudy	cloud free	<i>Méso-NH / MODIS</i>	cloudy	cloud free
cloudy	91%	67%	cloudy	79%	63%
cloud free	9%	33%	cloud free	21%	37%
<i>MM5 / MODIS</i>	cloudy	cloud free	<i>RACMO2 / MODIS</i>	cloudy	cloud free
cloudy	69%	54%	cloudy	85%	59%
cloud free	31%	46%	cloud free	15%	41%
<i>RCA / MODIS</i>	cloudy	cloud free			
cloudy	79%	79%			
cloud free	21%	21%			

Table 4

As Table 3 but for S2. The spatial distribution of the classes can be seen in Fig. 8.

<i>LM / MODIS</i>	cloudy	cloud free	<i>Més0-NH / MODIS</i>	cloudy	cloud free
cloudy	76%	39%	cloudy	80%	77%
cloud free	24%	61%	cloud free	20%	23%
<i>MM5 / MODIS</i>	cloudy	cloud free	<i>RACMO2 / MODIS</i>	cloudy	cloud free
cloudy	66%	23%	cloudy	83%	70%
cloud free	34%	77%	cloud free	17%	30%
<i>RCA / MODIS</i>	cloudy	cloud free			
cloudy	84%	93%			
cloud free	16%	7%			

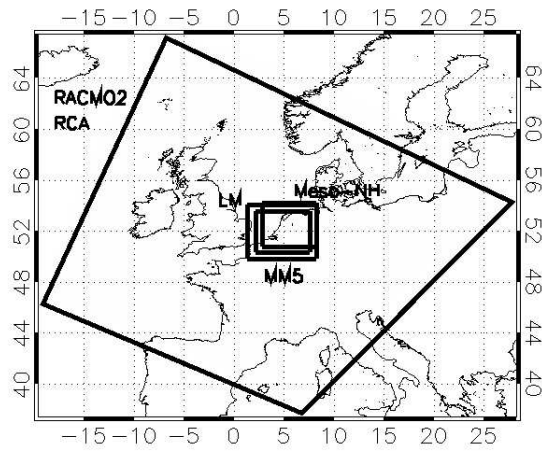


Fig. 1

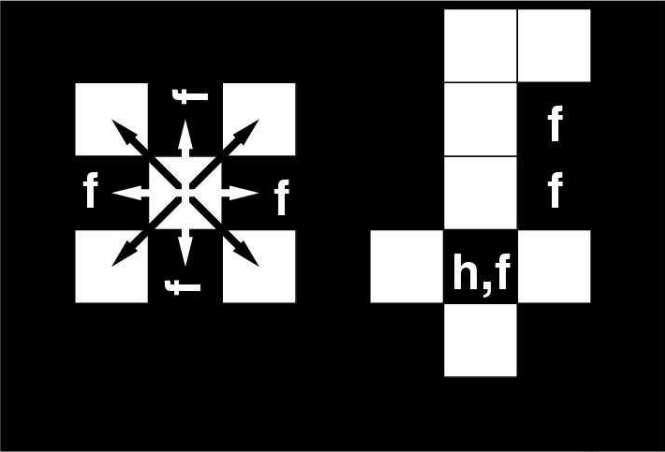


Fig. 2

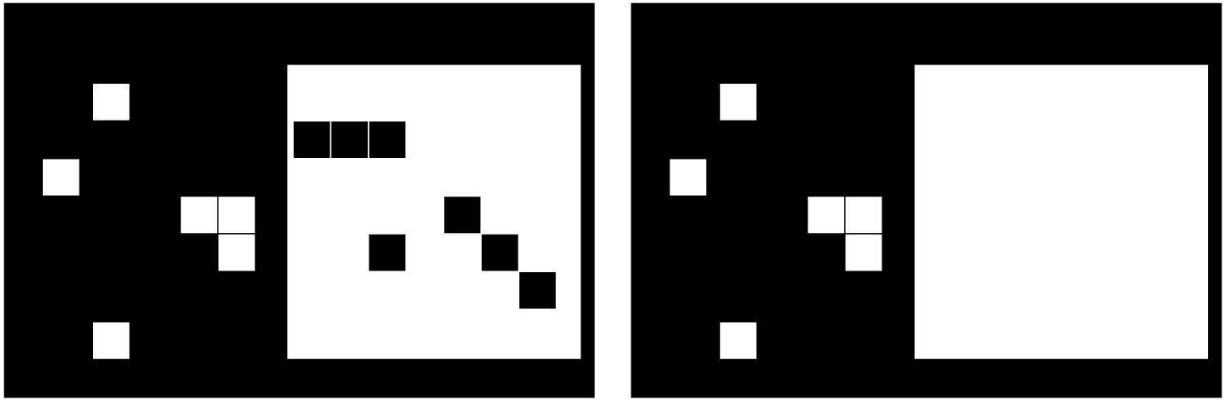


Fig. 3

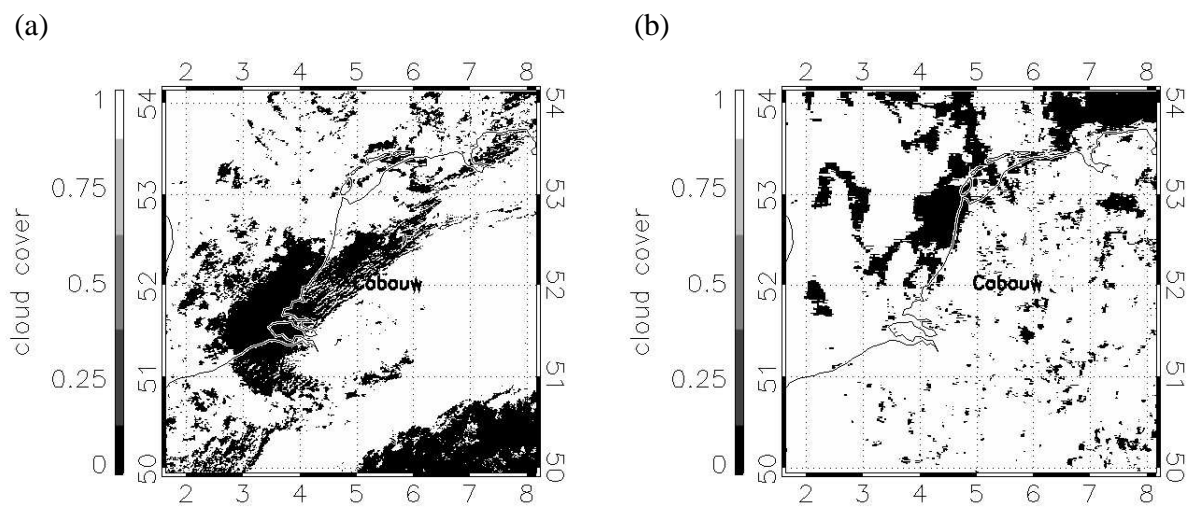


Fig. 4

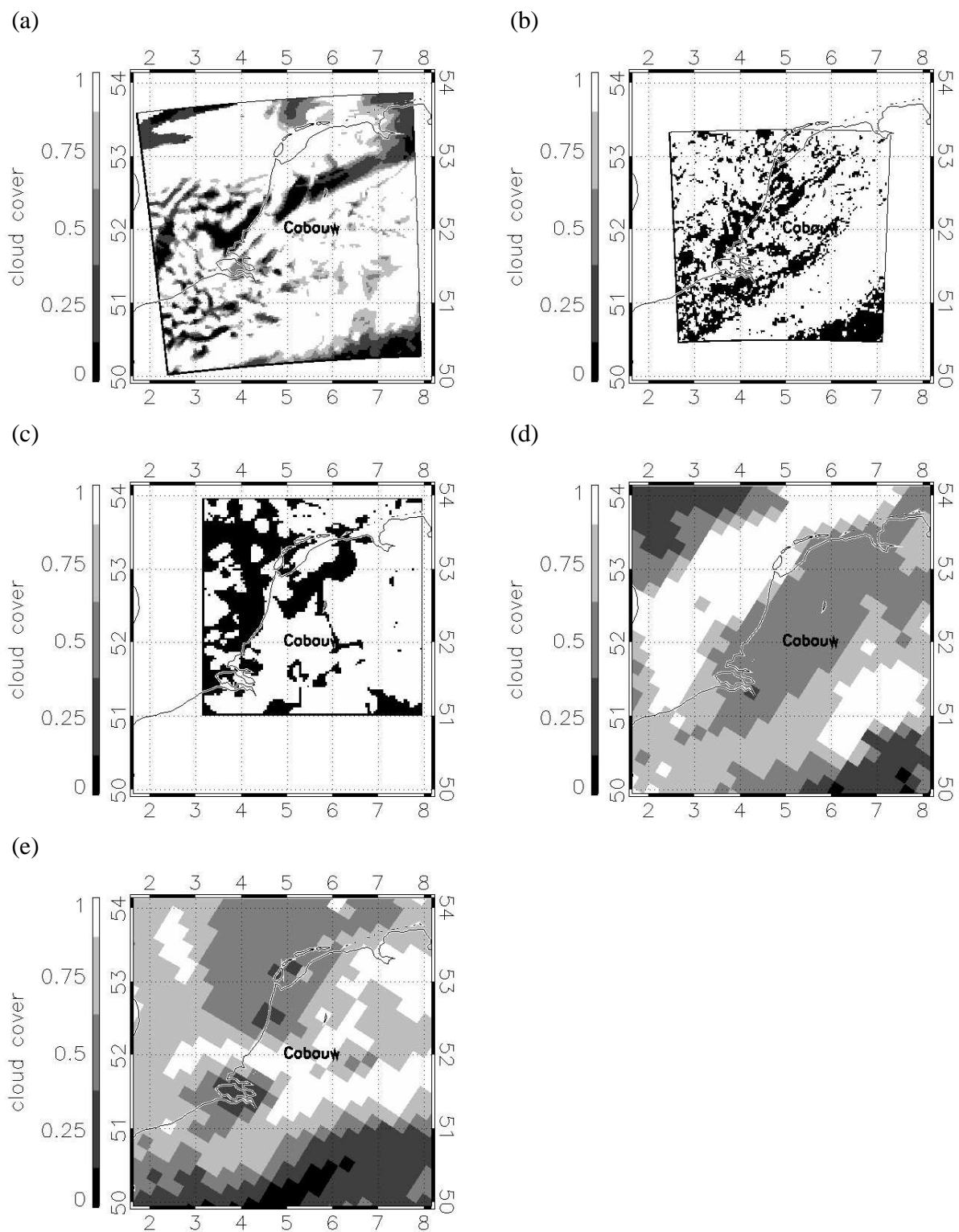


Fig. 5

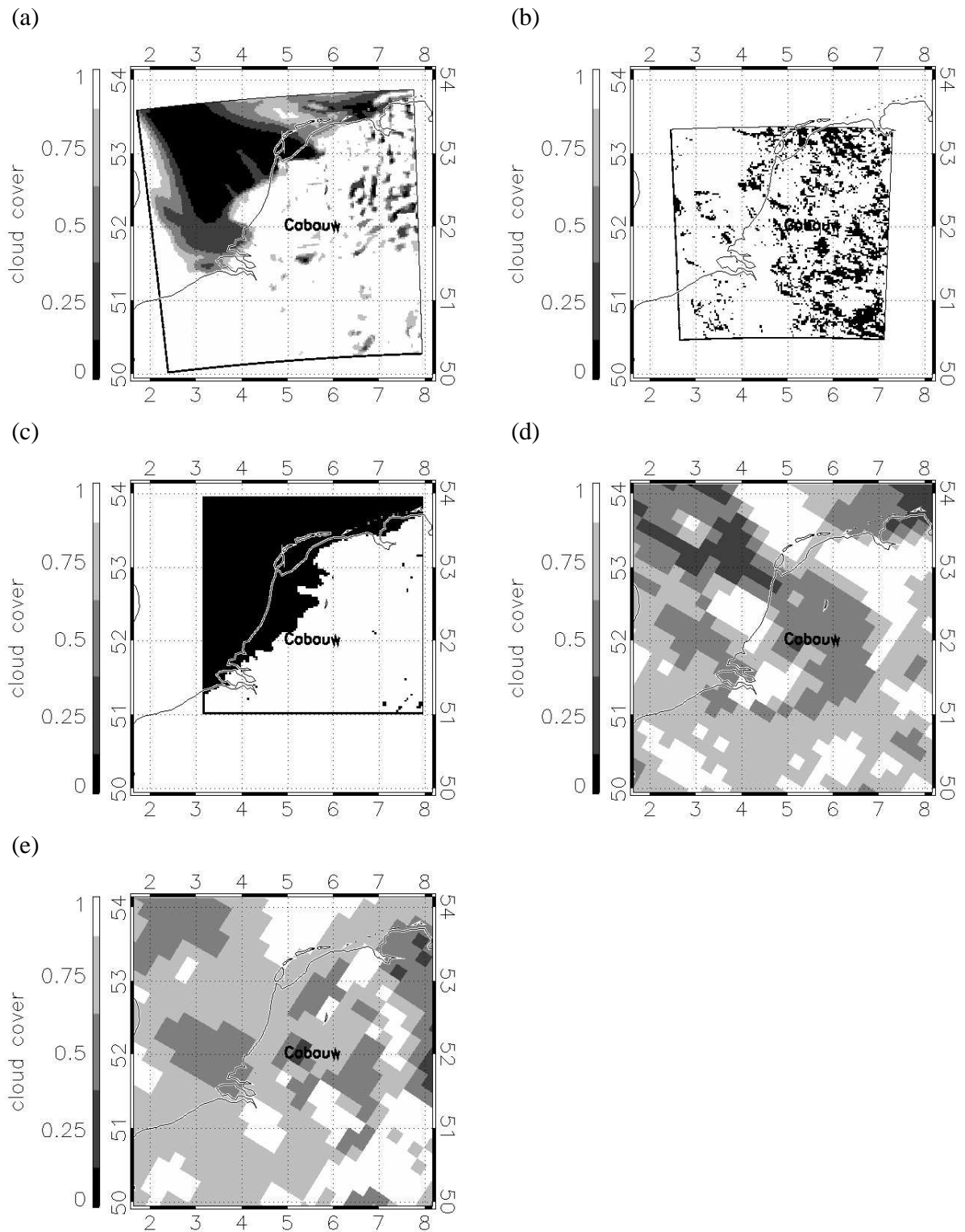


Fig. 6

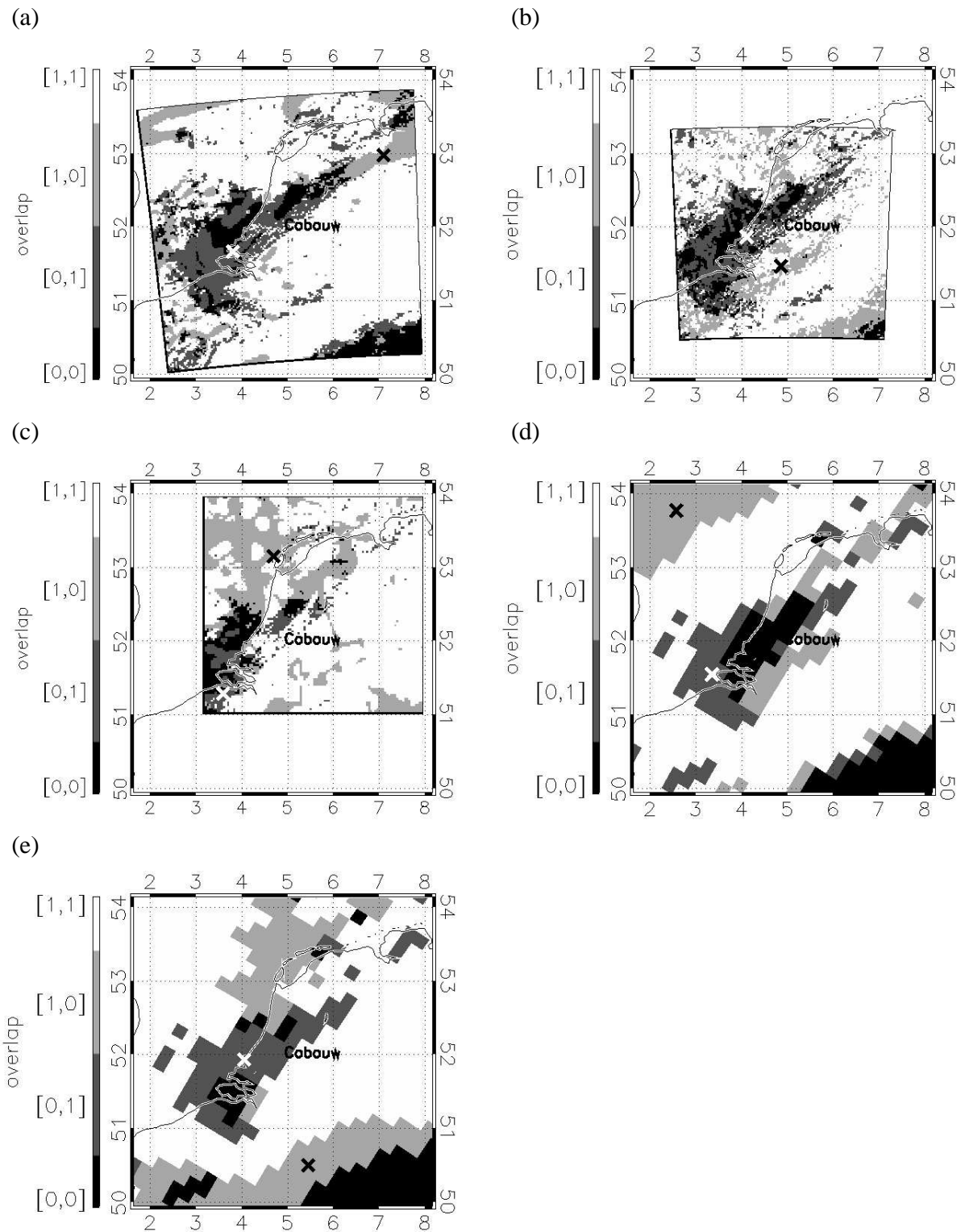


Fig. 7

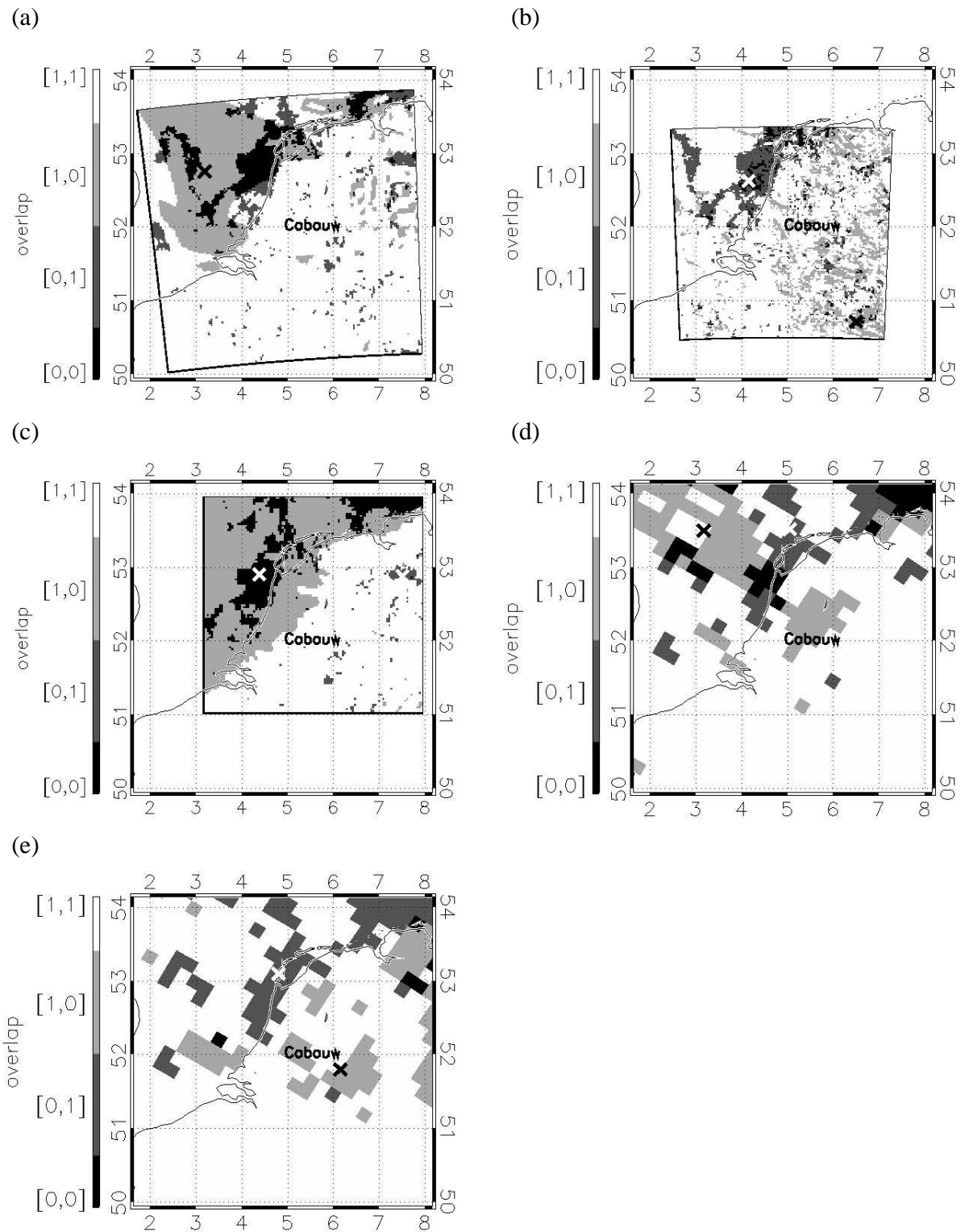


Fig. 8

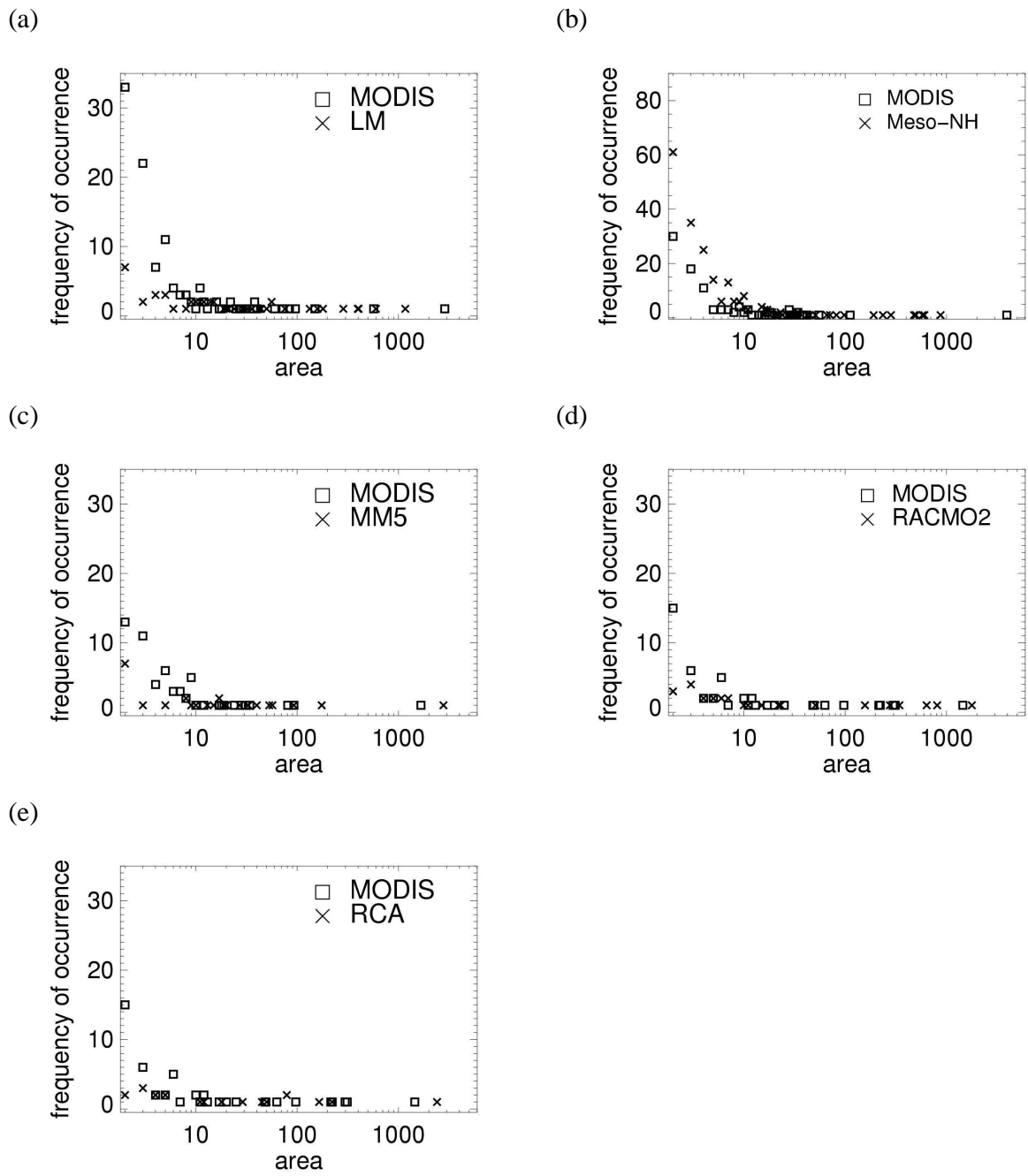
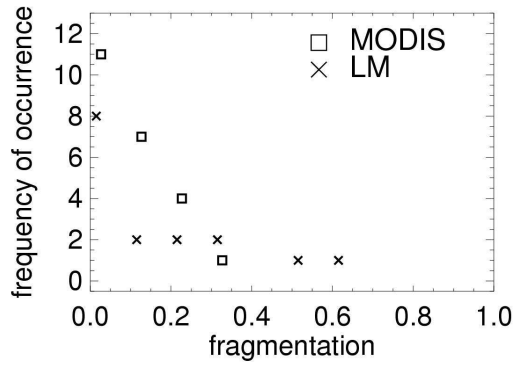
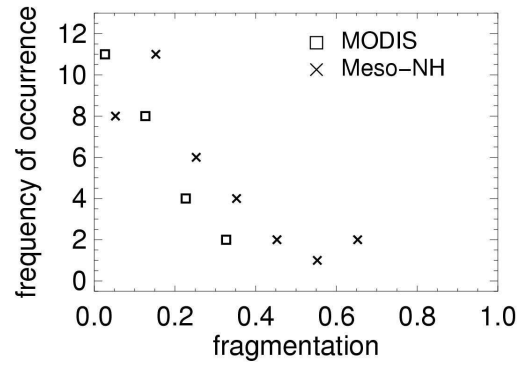


Fig. 9

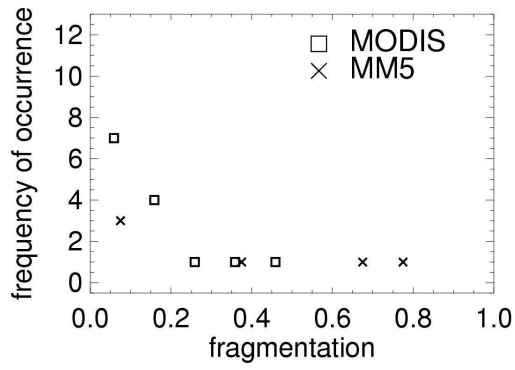
(a)



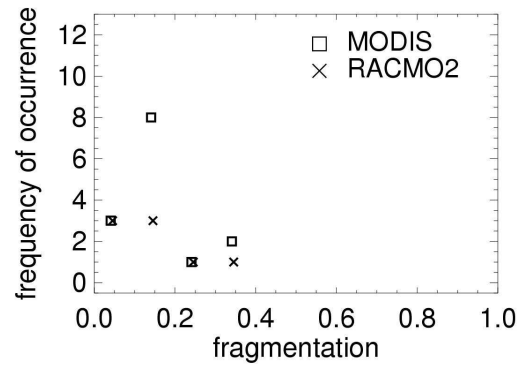
(b)



(c)



(d)



(e)

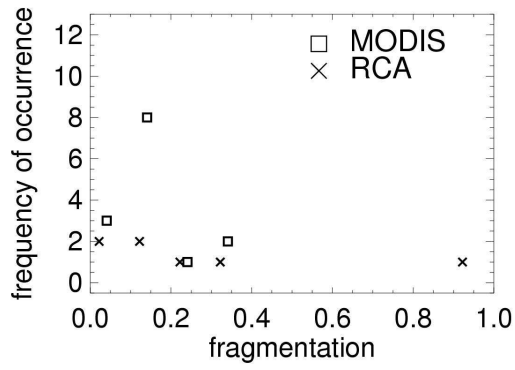


Fig. 10

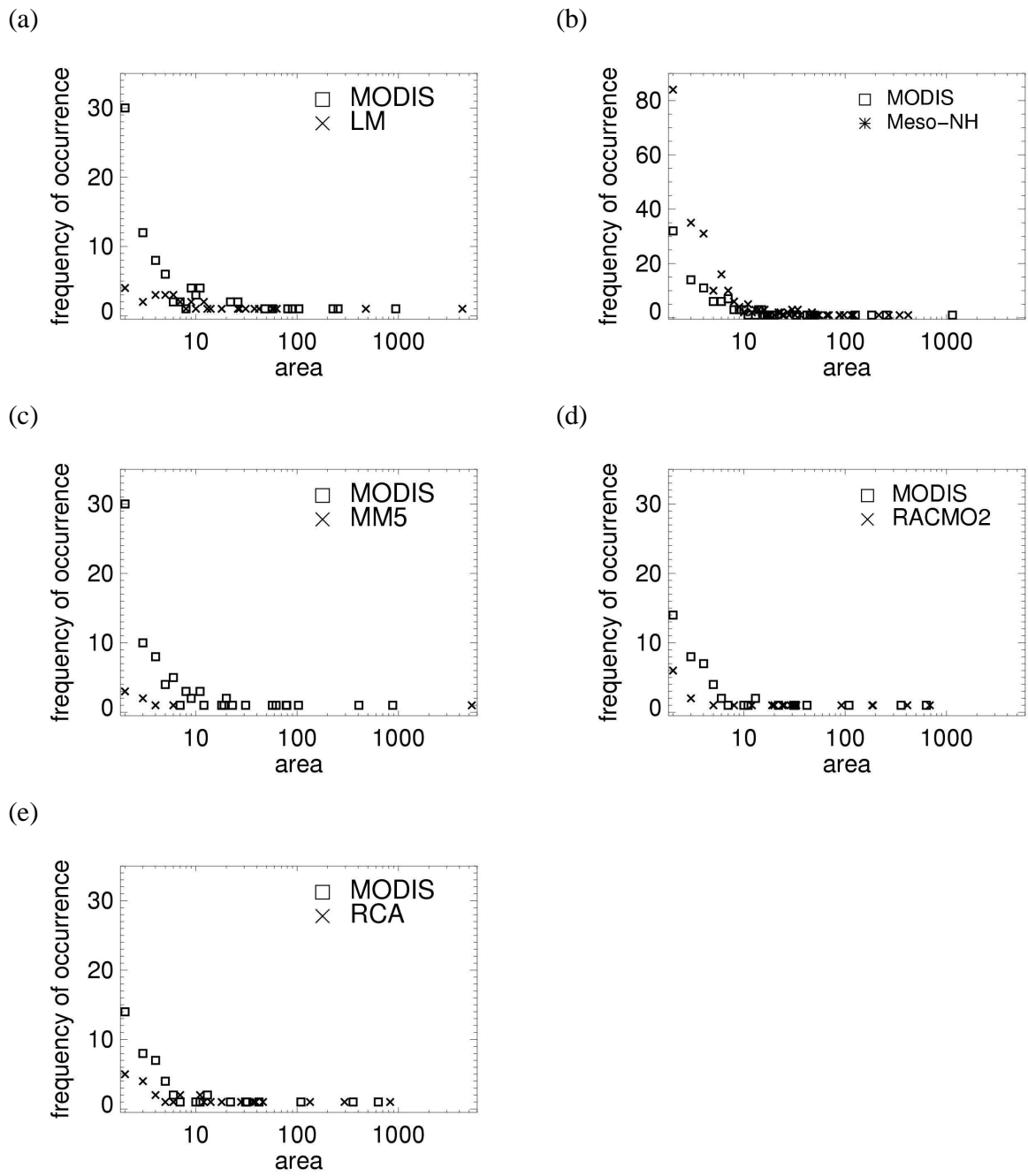


Fig. 11

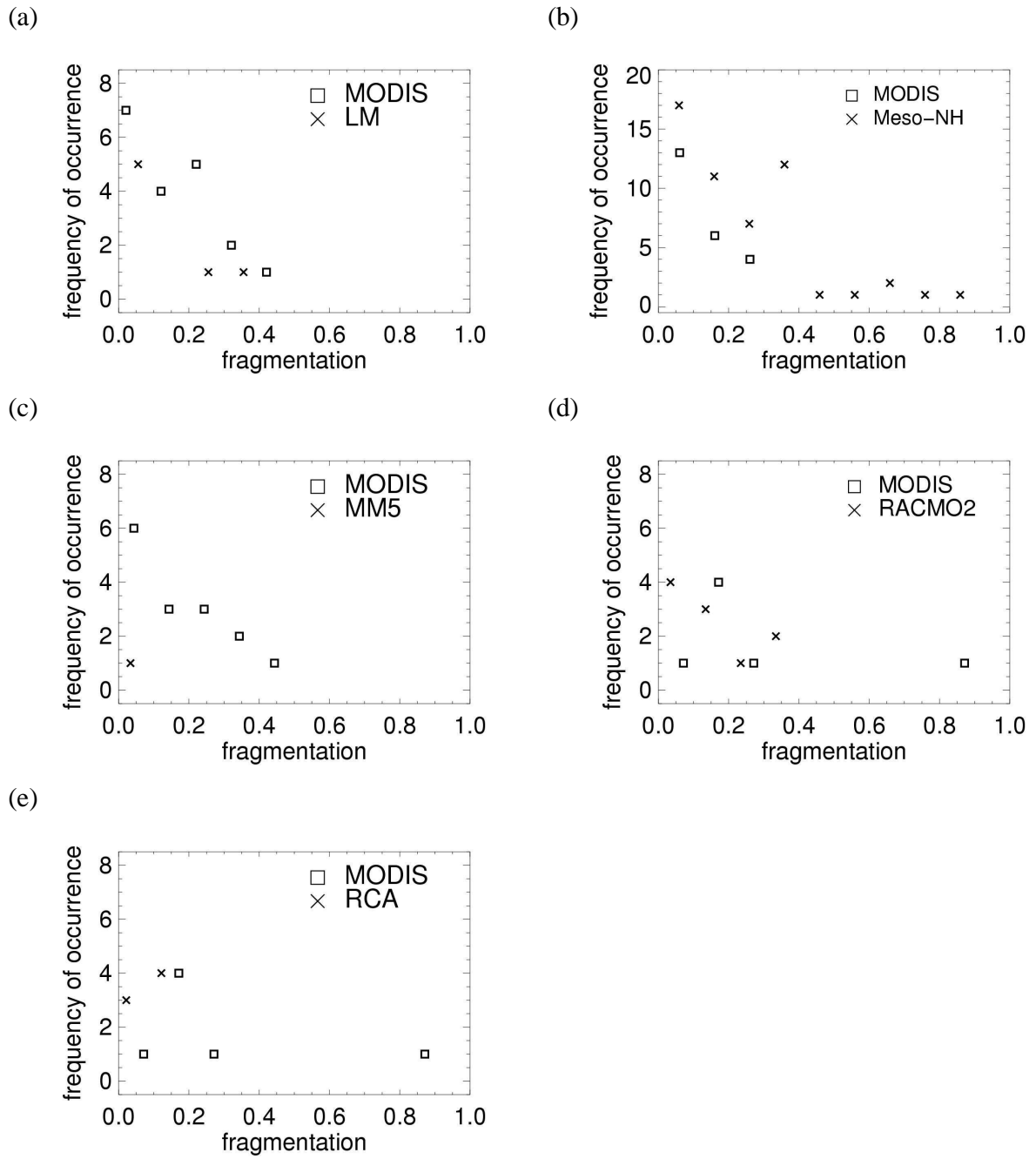


Fig. 12