Atmospheric
Measurement
Techniques

High-resolution satellite-based cloud detection for the analysis of land surface effects on boundary layer clouds

Julia Fuchs[1,2], **Hendrik Andersen**[1,2], **Jan Cermak**[1,2], **Eva Pauli**[1,2], **and Rob Roebeling**[3]

[1]Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research, Karlsruhe, Germany
[2]Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Karlsruhe, Germany
[3]European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany

Correspondence: Julia Fuchs (julia.fuchs@kit.edu)

Abstract. The observation of boundary layer clouds with high-resolution satellite data can provide comprehensive insights into spatiotemporal patterns of land-surface-driven modification of cloud occurrence, such as the diurnal variation of the occurrence of fog holes and cloud enhancements attributed to the impact of the urban heat island. High-resolution satellite-based cloud-masking approaches are often based on locally optimised thresholds that can be affected by the local surface reflectance, and they therefore introduce spatial biases in the detected cloud cover. In this study, geostationary satellite observations are used to develop and validate two high-resolution cloud-masking approaches for the region of Paris to show and improve applicability for analyses of urban effects on clouds. Firstly, the Local Empirical Cloud Detection Approach (LECDA) uses an optimised threshold to separate the distribution of visible reflectances into cloudy and clear sky for each individual pixel accounting for its locally specific brightness. Secondly, the Regional Empirical Cloud Detection Approach (RECDA) uses visible reflectance thresholds that are independent of surface reflection at the observed location. Validation against in-situ cloud fractions reveals that both approaches perform similarly, with a probability of detection (POD) of 0.77 and 0.69 for LECDA and RECDA, respectively. Results show that with the application of RECDA a decrease of cloud cover during typical fog or low-stratus conditions over the urban area of Paris for the month of November is likely a result of urban effects on cloud dissipation. While LECDA is representative for the widespread usage of locally optimised approaches, comparison against RECDA reveals that the cloud masks obtained from LECDA result in regional biases of ±5 % that are

most likely caused by the differences in surface reflectance in and around the urban areas of Paris. This makes the regional approach, RECDA, a more appropriate choice for the high-resolution satellite-based analysis of cloud cover modifications over different surface types and the interpretation of locally induced cloud processes. Further, this approach is potentially transferable to other regions and temporal scales for analysing long-term natural and anthropogenic impacts of land cover changes on clouds.

1 Introduction

The detection of continental boundary layer clouds based on satellite data has a long tradition and is essential for the analysis of the various ways that clouds interact with the land surface from the climate scale to the microscale. Different land cover types, such as vegetation and urban surfaces, exhibit distinct physical surface properties that influence the latent and sensible heat fluxes between the surface and the boundary layer and thus cloud development and spatial patterns (Shepherd, 2005; Collier, 2006; Varentsov et al., 2018).

Frequently, satellite-based analyses of the impact of land surface characteristics on boundary layer clouds are based on comparisons of cloud fraction and its spatial anomalies with respect to different land cover types (Teuling et al., 2017; Theeuwes et al., 2019; Pauli et al., 2022). Characteristic modifications of spatial patterns of cloudiness can emerge over urban areas. For example, an afternoon cloud cover enhancement has been found over Paris during summer (Theeuwes et al., 2019), whereas fog frequencies have been

observed to be lower over the urban areas of the Gangetic Plain during winter (Gautam and Singh, 2018). In this context, the urban heat island effect has been found to be a main driver of the boundary layer cloud enhancement over Paris (Theeuwes et al., 2019). In many cases, the mechanisms by which the land surface characteristics interact with the spatial patterns in cloudiness are not well understood (Zhong et al., 2017; Fan et al., 2016; Liang et al., 2018; Shepherd, 2005; Collier, 2006; Han et al., 2014).

Satellite data are ideally suited to map patterns of cloudiness and facilitate the analysis of land surface–atmosphere interactions at various scales. The High Resolution Visible (HRV) channel of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor onboard Meteosat Second Generation (MSG) is particularly useful in this context due to its high temporal ($\sim$ 15 min) and spatial resolutions (1 km at nadir) (Klüser et al., 2008; Deneke and Roebeling, 2010; Derrien et al., 2010; Henken et al., 2011). High-resolution cloud masks enable the detection and analyses of modifications to spatial patterns of cloudiness induced by small-scale (effective spatial resolution of HRV in this region $\sim$ 2 km) features. While the higher spatial resolution is a clear advantage of the HRV channel, cloud detection from the single spectral HRV channel is hampered by missing multi-spectral information that is available at coarser resolutions (e.g. 3 km at nadir for SEVIRI). This is shown by the large number of cloud detection algorithms that are based on multi-spectral thresholding tests to separate clear sky from different cloud types (Ackerman et al., 1998; Rossow and Garder, 1993; Saunders and Kriebel, 1988; Stowe et al., 1999; Di Vittorio and Emery, 2002; Chen et al., 2003; Hutchison et al., 2005; Cermak, 2006; Yang et al., 2007; Bley and Deneke, 2013; Andersen and Cermak, 2018).

HRV-based cloud detection methods for deriving cloud masks are rare and challenging due to the limitation of observations from a single visible channel (Schulz et al., 2012; Nilo et al., 2018). Cloud masks obtained from a single channel are often based on histograms and seek to separate the clear-sky and cloudy components of the reflectivity distribution. The separation of the distribution is typically based on a threshold that is determined by a histogram-based minima approach in a representative data set of a specific surface type and geographic region (Minnis and Harrison, 1984; Ipe et al., 2003; Bley and Deneke, 2013; Cermak, 2006; Cermak and Bendix, 2008). For the empirical determination of the clear-sky reflectivity, the frequent occurrence of atmospheric and non-atmospheric features can be challenging, as aerosols, cloud shadows, thin cirrus clouds and fresh snow affect the retrieved reflectances (Yang et al., 2007; Ipe et al., 2003; Matthews and Rossow, 1987). In general, thresholds determined by such techniques are temporally dynamic to account for the diurnal and seasonal variations of the solar zenith angle (Yang et al., 2007) and the vegetation period (Teuling et al., 2017; Theeuwes et al., 2019). The main problem for cloud detection is the influence of the land surface

signal on the clear-sky determination that can lead to severe misinterpretation of cloud occurrence over different surface types. For example, over bright surfaces a reduced contrast between clear-sky reflection and cloudy-sky reflection can lead to an under-detection of clouds (EUMETSAT, 2019). To account for location-specific differences in surface albedo and to reduce the local bias, thresholds determined over land are frequently locally optimised. Teuling et al. (2017) suggested an empirical approach that flags an observation as cloudy when reflectivity exceeds a locally (per pixel) and temporally (per hour, 10 d period) adjusted clear-sky climatological value. Such locally adjusted and empirically based thresholds have a low bias in total but are still challenged by a degradation of detection accuracy depending on the local surface brightness (Bley and Deneke, 2013). This is caused by locally optimised approaches that seek to find the optimal threshold to distinguish clouds and clear sky and thus have to deal with a surface-dependent bias (see explanation in Sects. 3.2 and 1b). For analyses of land surface impacts on clouds, however, it is crucial that the satellite-based detection of cloud occurrence over different land cover types is independent of variations in surface properties, like spectral albedo. With such an application-focused algorithm there is a great potential to further improve the understanding of cloud formation and dissipation processes with respect to regional characteristics.

In this study, two approaches to detect boundary layer clouds on the basis of SEVIRI HRV data are presented and compared for the analysis of cloud cover changes over different surface types: the locally optimised cloud detection scheme of the Local Empirical Cloud Detection Approach (LECDA) and the Regional Empirical Cloud Detection Approach (RECDA) that is more robust to regional variations in surface reflectivity. Both approaches are processed for November of a multi-year period over a large region of Paris and its surrounding area. During this month, fog occurs frequently and can be expected to be affected by the urban region of Paris (Gautam and Singh, 2018). Results are validated with measurements of a CloudNet station in the southwest of Paris.

In Sect. 2, the data used are described. The methods to derive LECDA and RECDA and their validation with in situ CloudNet data are explained in Sect. 3. Section 4 presents the results, a discussion of the comparison and the validation of the two cloud-masking approaches, and a meteorological case using RECDA. Section 5 contains the conclusions.

2   Data

2.1   SEVIRI satellite data

The main part of this study, including the generation of cloud masks, is based on the analysis of SEVIRI HRV data, which have a spatial resolution of 1 km at nadir and cover a broad-

band range from 0.4 to 1.1 µm. Together with the remaining three solar and eight thermal SEVIRI channels with a spatial sampling distance of 3 km at nadir, the Earth is observed in a repeat cycle of 15 min covering the full disc for the low-resolution channels and half of the full disc for the HRV channels. It must be noted that spatial resolutions of the low- and high-resolution channels are effectively lower than the given sampling distance due to the satellite viewing geometry and the fact that the reflectance field is oversampled by a factor of 1.6 (Deneke and Roebeling, 2010; Schmetz et al., 2002, MSG level 1.5 image data format description). For the detection and removal of snow and ice clouds, channel reflectances of the visible channels (0.6 and 1.6 µm) and brightness temperatures of the infrared channels (8.7, 10.8 and 12.0 µm) are used, respectively (see Sect. 3.1). In total, the time period of the data set spans the month of November from 2004 to 2019 (08:00–16:00 UTC), resulting in ∼ 14 400 time steps. The study region is the urban area of Paris and the adjacent rural areas extending from 48.0 to 49.6° N and from 1.6 to 3.0° E (Fig. 4b), where urban cloud modifications have already been observed and are considered relevant for the urban climate of Paris during summer (Theeuwes et al., 2019). The urban area of Paris further serves as an ideal test bed as it is surrounded by relatively homogeneous and flat terrain with altitude ranges of ∼ 200 m (see Fig. 6b). The month of November was chosen to investigate a possible reduction in fog clouds over the urban area of Paris as found in other regions during winter (Gautam and Singh, 2018). Further, with the selection of 1 month, seasonal variations of land cover, e.g. due to the vegetation period, can be neglected and long-term variability of the surface reflectance can assumed to be small.

2.2 CloudNet data

The cloud masks are validated against the calibrated Cloud-Net classification product (CF-1.0, level 2) for November (2015 to 2019) at Palaiseau (48.718° N–2.202° E) located ∼ 25 km to the southwest of Paris at 156 m a.s.l. (Illingworth et al., 2007). The location of the CloudNet station is shown in Fig. 2. Ground-based remote sensing instrumentation, including a ceilometer, a Doppler cloud radar, a microwave radiometer from the atmospheric observatory SIRTA (Site Instrumental de Recherche par Télédétection Atmosphérique, Haeffelin et al., 2005) and model outputs of the ECMWF Integrated Forecast System (IFS), serve as input for the Cloud-Net target classification algorithms described in ACTRIS Deliverable WP5 D5.5 (M24). As a result, a classification into 11 classes distinguishes between ice and water clouds, precipitation, aerosols, insects, clear sky, and combinations thereof (See Table A1 in Appendix A) is obtained in a 30 s time interval and for 719 heights above sea level with a vertical resolution of 25 m (168.5–18 118.5 m).

2.3 Reanalysis data

To highlight the applicability of the novel cloud mask data set generated by the regional approach to study surface-driven modifications of fog cloud cover, ERA5 data are used to filter for atmospheric conditions under which these clouds typically occur. ERA5 hourly data on single levels from 1979 to present of the European Centre for Medium-Range Weather Forecasts (ECMWF; Hersbach et al., 2018) are averaged over the study region, and SEVIRI scenes are selected only if all of the following criteria are met: 10 m wind speed that is $< 3\,\mathrm{m\,s^{-1}}$ (U/V wind component), boundary layer height (BLH) that is $< 300\,\mathrm{m}$ and mean sea level pressure (MSL) that is $> 1020\,\mathrm{hPa}$. These are boundary layer conditions that characterise and facilitate the formation of fog and low-strata cloud.

2.4 CORINE land cover

For comparisons between cloud fraction anomalies over different land cover types, CORINE land cover (CLC) level 3 data for 2012 (version 2020_20u1) at 100 m spatial resolution provided by the Copernicus Land Monitoring Service are used (European Environment Agency, EEA, 2021). The land use in 2012 only shows minor differences compared to the available land cover classifications in 2006 and 2018 and is chosen here to represent the period 2004–2019. CLC 2012 data are resampled to the HRV spatial resolution by using the most frequent land cover class within a HRV pixel (Fig. 4a). Only relevant land cover classes, including forests, continuous or discontinuous urban fabric, arable land, and pastures, are considered for a more detailed analysis.

2.5 Digital elevation model

The European Digital Elevation Model (EU-DEM; version 1.1) is used to compare changes in altitude of small-scale features of cloud fraction anomalies (Fig. 6b). The data set has a spatial resolution of 25 m and is provided by the Copernicus Land Monitoring Service (EEA, 2017; Tøttrup, 2014).

3 Methods

3.1 Snow and ice filter

The preprocessing of the HRV data includes the elimination of ice clouds and snow in order to retain only liquid and mixed clouds for the data analysis. This is done by using threshold tests on the SEVIRI channels at 0.6, 1.6, 8.7, 10.8 and 12 µm and collocating the flagged pixels with the associated HRV pixels. In this way, data that are not relevant for the analysis of urban cloud modifications are excluded from the data set.

Snow is detected by calculating the Normalised Difference Snow Index (NDSI) based on the 0.6 and 1.6 μm channels:

$$\text{NDSI} = \frac{r_{0.6} - r_{1.6}}{r_{0.6} + r_{1.6}}, \tag{1}$$

where r_λ is the reflectance at wavelength λ (μm) (Dozier and Painter, 2004; Cermak, 2006). Tests done for specific scenes revealed a threshold of 0.3 to be suitable to detect snow. Pixels that feature an NDSI above this value are excluded from the data set. This approach has been widely and operationally applied to different satellite data, including MODIS and Sentinel-2 satellite sensors (Hall et al., 2001; Richter et al., 2012).

A second filter is applied to exclude ice clouds from the data set with a brightness temperature below 263 K in the 10.8 μm channel. To exclude potential remaining ice clouds, the difference between the 8.7 and 12 μm channels is used in a phase test to retain only liquid clouds. Ice absorbs much more strongly than liquid water between 10 and 13 μm, while both show a similar absorption pattern from 8 to 10 μm (Strabala et al., 1994; Cermak, 2006). The brightness difference should thus be smaller for ice clouds and is implemented as in Westerhuis et al. (2020) with the fog and low-stratus confidence level CL_{FLS}:

$$\text{CL}_{\text{FLS}} = \frac{T_{\text{LC}} - (T_{12\,\mu m} - T_{8.7\,\mu m}) - T_{\text{CCR}}}{-2T_{\text{CCR}}}, \tag{2}$$

where $T_{\text{LC}} = 1.8\,\text{K}$ is the threshold for liquid clouds and $T_{\text{CCR}} = 1\,\text{K}$ is the cloud confidence range. Details of this approach can be found in Westerhuis et al. (2020). Pixel values with a CL_{FLS} below 0 are removed from the data set to retain only low-level liquid clouds.

3.2 Two approaches to delineate clouds and clear sky

3.2.1 Gaussian mixture model

Based on this preprocessed data set the Local Empirical Cloud Detection Approach and the Regional Empirical Cloud Detection Approach is proposed to delineate clouds from clear sky resulting in two different cloud masks (CMloc and CMreg). The ultimate goal of RECDA is to derive spatially unbiased cloud fraction anomalies over different land cover types, while the goal of LECDA is a pixel-by-pixel cloud detection with minimum bias.

Both approaches are relatively simple and nearly independent of other channel information. The thresholds are determined empirically based on occurrence frequencies of HRV counts in solar zenith angle (SZA) bins as shown conceptually in Fig. 1a for a single SZA bin and pixel. The first peak of the distribution presents clear-sky conditions (CS), while the second refers to cloudy (CL) conditions.

In both approaches the CS composite is obtained by applying a Gaussian mixture model (GMM) to the HRV reflectance histograms of each SZA bin and each pixel within the study region. A GMM is a probabilistic model that assumes that the data are generated from a mixture of a finite number (here this number is two: CS and CL in Fig. 1a) of Gaussian probability distributions with unknown parameters. This method is categorised as an unsupervised learning algorithm that works in a similar way as the distance-based k-means cluster method. Instead of a distance-based model, the algorithm finds the clusters based on Gaussian probability distributions. Based on the expectation minimisation algorithm, conditional expectations of the complete log likelihood are calculated, and estimates (mean and covariance) are updated iteratively until convergence to a local optimum is reached (Jain et al., 2000). While this method is limited to near-Gaussian distributions of a representative data set (excluding high SZA bins, where a bimodal distribution of clear sky and cloud is not always existent) and a sufficient amount of data, it has great potential beyond the simple usage of a binary cloud mask as it provides probabilities (Sci-kit learn implementation, Pedregosa et al., 2011).

In this setup, the GMM characteristics of the CS component are exploited to dynamically (pixel-based per SZA bin) derive thresholds. In the following, the computation of a local T_{loc} (1) and a regional threshold T_{reg} (2) is described based on the maximum bin count and standard deviation of the GMM CS component.

3.2.2 Local Empirical Cloud Detection Approach (LECDA)

In this approach a local HRV threshold T_{loc} is obtained based on the CS probability distribution (first GMM component) per SZA slot (five SZA bins 67–77; > 1000 data points) and per pixel as follows. First, the local clear-sky reflectance CS_{loc} (Fig. 1a), which is determined by the reflectance value using the maximum count per histogram bin (maxbin) with a bin width of 0.5 within the mean $\pm 2\sigma$ range of the CS GMM maxbin approach, is preferred over taking the mean of the first GMM component as it is representative for CS. Second, T_{loc} is computed as the sum of CS_{loc} (per pixel) and the regional median of all local 3σ of the CS GMM component. The regional median is chosen to increase the robustness of the derived T_{loc} against a sub-optimal local fit of the GMMs due to atmospheric noise, such as aerosols. In general, this method is assumed to be less contaminated by the effect of cloud shadows, aerosols and thin cirrus clouds compared to similar clear-sky minimum approaches as it is based on the maximum bin count of the GMM clear-sky distribution (Bley and Deneke, 2013; EUMETSAT, 2019). However, the impact of these atmospheric features on the results cannot be entirely excluded and may distort the clear-sky reflectance.

$$\text{CS}_{\text{loc}} = \text{maxbin} \in \left[\text{mean} + 2\sigma_{\text{CSGMM}}, \text{mean} - 2\sigma_{\text{CSGMM}}\right] \tag{3}$$

$$T_{\text{loc}} = \text{CS}_{\text{loc}} + \text{MED}_{\text{reg}}(3\sigma_{\text{CSGMM}}) \tag{4}$$

Thresholds obtained by LECDA as in Eq. (4) are used in, for example, Teuling et al. (2017) and present a straight-
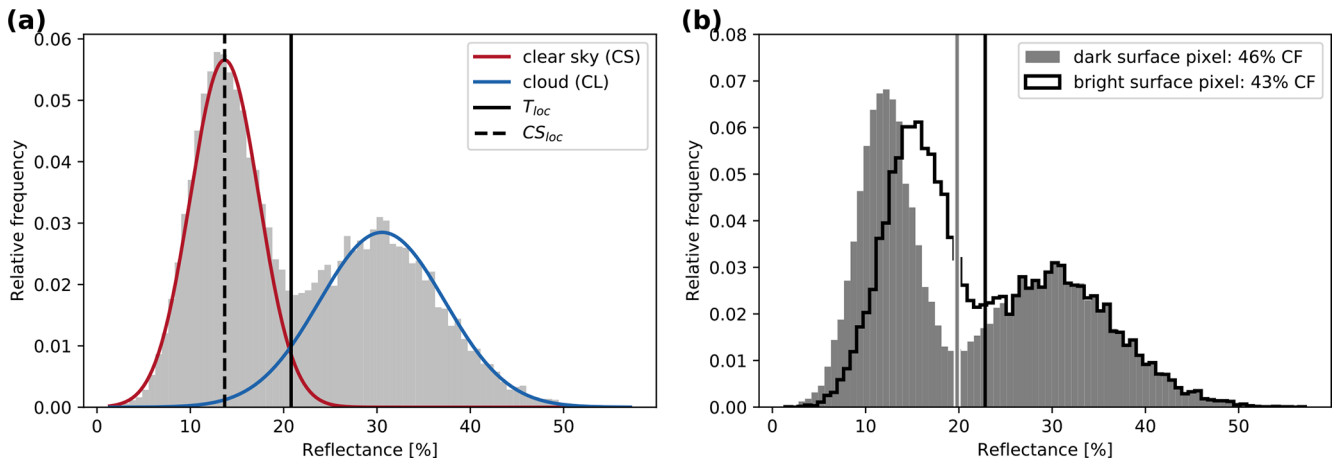
Figure 1. (a) Conceptional distribution of a bimodal distribution of HRV pixel counts for one SZA bin and one pixel within the study region during November 2004–2019. The first and second components of the Gaussian mixture model present clear (red) and cloudy conditions (blue) with the local clear-sky reflectance CS_{loc} (vertical dashed black line) and a potential clear-cloud threshold T_{loc} (vertical black line). Clear-sky reflectance is defined as the maximum count per histogram bin (maxbin) of the CS GMM component. The histogram minimum would be a potential threshold to separate the near-Gaussian distributions of clear and cloudy pixels for this time and location. Panel **(b)** is the same as **(a)** over a bright and a dark surface with the corresponding cloud fraction (CF). Both figures are created with three randomly generated Gaussian distributions: a distribution for the cloudy conditions and two distributions for surfaces of different brightness (Sect. 3.2), showing the surface-brightness-driven shift in the threshold that leads to different estimates of cloud fraction in this artificial example.

forward method to approximate the "real" cloud fraction. In their study a similar local cloud mask approach is proposed, where a constant threshold (10 counts) is added to an empirically determined clear-sky value for a given pixel based on a cumulative distribution function of reflectivity measurements. However, such locally optimised approaches including LECDA are dependent on variations in surface albedo due to different land cover types as a constant value is added to the clear-sky reflectance shown in Fig. 2 and Eq. (4). A brighter clear-sky surface pixel within the urban area of Paris will shift the clear-sky distribution to higher reflectance and thus clear-cloud thresholds assuming a constant cloud distribution as schematically illustrated using synthetic data in Fig. 1b. Due to the dynamic threshold method (pixel-based per SZA bin) that seeks to distinguish bright clear-sky surface pixels from clouds in a narrower bimodal distribution, a higher threshold is set in the case of brighter pixels compared to darker pixels. As a direct consequence, the application of LECDA for the retrieval of cloud fraction can potentially result in an underestimation of clouds over bright surfaces when compared to clouds over dark surfaces (see the legend of Fig. 1b for an example) given a constant cloud distribution over both surfaces.

As comparability of cloud detection over different surface albedos is not assured by LECDA, and to address the potentially resulting bias, RECDA is developed to use one regional threshold for all pixels of the study area (Eq. 5).

3.2.3 Regional Empirical Cloud Detection Approach (RECDA)

In this approach T_{reg} is obtained by setting the maximum of all T_{loc} as a "static" threshold that is applied to all pixels in the study area per SZA.

$$T_{reg} = \max(T_{loc}) \qquad (5)$$

RECDA is thought to result in a slight underestimation of clouds for all pixels as it sets the reflectance threshold to a higher reflectance value. The advantage compared to the LECDA is that RECDA is independent of different surface albedo values (Fig. 2) and is therefore better suited to the comparison of regional cloud fraction anomalies over different land surfaces. The influence of the surface signal during the prevalence of thin liquid clouds may affect the cloud distribution but is assumed to be minor compared to the dependence of LECDA on the surface signal (see Sect. 4.2).

3.3 Local evaluation using CloudNet data

CloudNet data are used to evaluate and compare the local performance of both approaches. In order to aggregate the CloudNet data to the temporal resolution of SEVIRI (15 min interval), CloudNet cloud fraction is calculated for each time step based on a 1 h time window centred around the corresponding SEVIRI time slot. This time window was chosen according to findings by Deneke et al. (2009) that compared SEVIRI satellite pixel (3 km × 6 km) with ground-based radiometer measurements. A correction factor of 11 min was added to the SEVIRI time slot to match the SEVIRI nominal
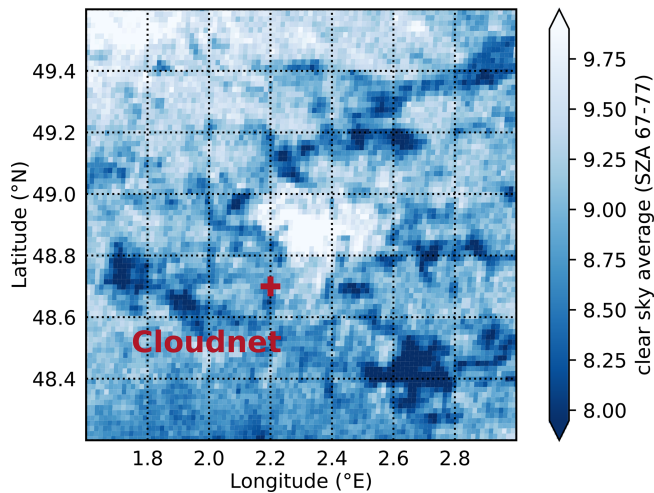
Figure 2. Average (SZA bins 67–77) of clear-sky reflectance CS_{loc} obtained by Eq. (3). The location of the CloudNet station at Palaiseau used for validation is marked in red.

time with the actual scanning time over the validation site. Its calculation is based on the time SEVIRI needs per revolution (0.6 s) to scan one line (three image lines at a time) from south to north, accounting for the acquisition start at 81° S, the spreading distance of the ground resolution in the S–N direction and the number of scan lines (MSG level 1.5 image data format description). The W–E scan time of 30 ms is neglected. Details of the computation for the low-resolution channels can be found in Kim et al. (2020) with the same nominal time as the HRV channel.

All CloudNet target classifications per time window are considered relevant for the validation of the cloud mask. To create a CloudNet cloud fraction as ground truth, classes 1 to 7 are assigned as cloudy, while 0 and 8 to 10 are assigned as clear sky (see Table A1; Veefkind et al., 2016). If one of the 719 vertical target layers at one specific time step is classified as cloud, the observation at this time step is flagged as cloud. If the resulting CloudNet cloud fraction of the defined time window is above 0.5, the CloudNet flag of the corresponding SEVIRI time is assigned as cloud (1) and is otherwise assigned as clear sky (0). This selected time window will ensure that only persistent cloud observations are matched with the SEVIRI cloud mask.

The temporally aggregated CloudNet cloud mask is compared to a spatial aggregate of 3×3 HRV cloud mask pixels ($\sim 3\,km \times 6\,km$) centred around the CloudNet validation station. Due to the oblique viewing angle of SEVIRI, a parallax correction is required, resulting in a horizontally displacement of the pixel matrix surrounding the validation station by $\sim 2.6\,km$ (two pixels) to the north (Greuell and Roebeling, 2009; Schutgens and Roebeling, 2009). If more than four out of the nine matrix pixels are cloudy, the matrix is aggregated to one value assigned as cloud (1) and is otherwise assigned as clear sky (0). This matrix approach was chosen in order

to fully capture clouds traversing the validation sight and to reduce the effect of cloud inhomogeneity and partial cloud cover (Deneke et al., 2005).

4   Results and discussion

4.1   Comparison of the two cloud-masking approaches

Figure 3 shows cloud fractions and their regional anomalies from the local and regional empirical cloud detection approaches, CFloc and CFreg. CFloc and CFreg are defined as the number of cloudy pixels divided by the total number of pixels, and their regional anomaly (CF$_{loc}$anomaly, CF$_{reg}$anomaly) is defined as the difference between the cloud fraction of the pixel and the average cloud fraction of the region. The spatial patterns of CF$_{loc}$ and CF$_{reg}$ in Fig. 3a and c are similar and show a high cloud fraction and hence a positive anomaly in Fig. 3b and d in a triangular shape east and northeast of Paris where the river valleys of the Oise, Marne and Seine are located. While CF$_{loc}$ shows generally higher cloud fractions and a slightly larger anomaly range with more small-scale features, CF$_{reg}$ spatial patterns are less distinct and more smooth. As expected, due to the algorithm differences (see Sect. 3.2), CF$_{reg}$ is lower (on average by 7 %) than CF$_{loc}$.

Differences between the two cloud-masking approaches are shown in Fig. 4a by mapping the difference of their respective cloud fraction anomaly patterns (CF$_{loc}$anomaly-CF$_{reg}$anomaly). The differences are most prominent over the urban region of Paris and over forests (cf. Fig. 4b) and are clearly connected with the clear-sky reflectance of these different land cover types (cf. Fig. 2). The comparison of the respective cloud fraction anomalies reveals a relative underestimation of CF$_{loc}$ by more than 4 % over the relatively bright urban area of Paris, while it leads to a relative overestimation of clouds over some of the relatively dark forest areas. The general dependence of LECDA on the surface reflectance is shown in Fig. 5a, where the difference between LECDA and RECDA is a function of clear0sky surface reflectance. A total of 75 % of the variability in the difference between RECDA and LECDA can be explained by the surface reflectance. It is assumed that the dependence on the surface reflectance is mostly attributed to LECDA, while a small portion may be attributed to RECDA in conditions of thin liquid clouds. Grouping these cloud fraction anomaly differences by land cover types confirms these observations and shows that the relative biases over the continuous urban area of Paris exceed the biases over the forest region (Fig. 5b). The pasture land cover type present in the northwestern parts of the region also shows a relative underestimation of the local approach; however, local differences from the surrounding land cover types are not clearly visible. The contribution of land cover variability within a HRV pixel (see resampling of land cover
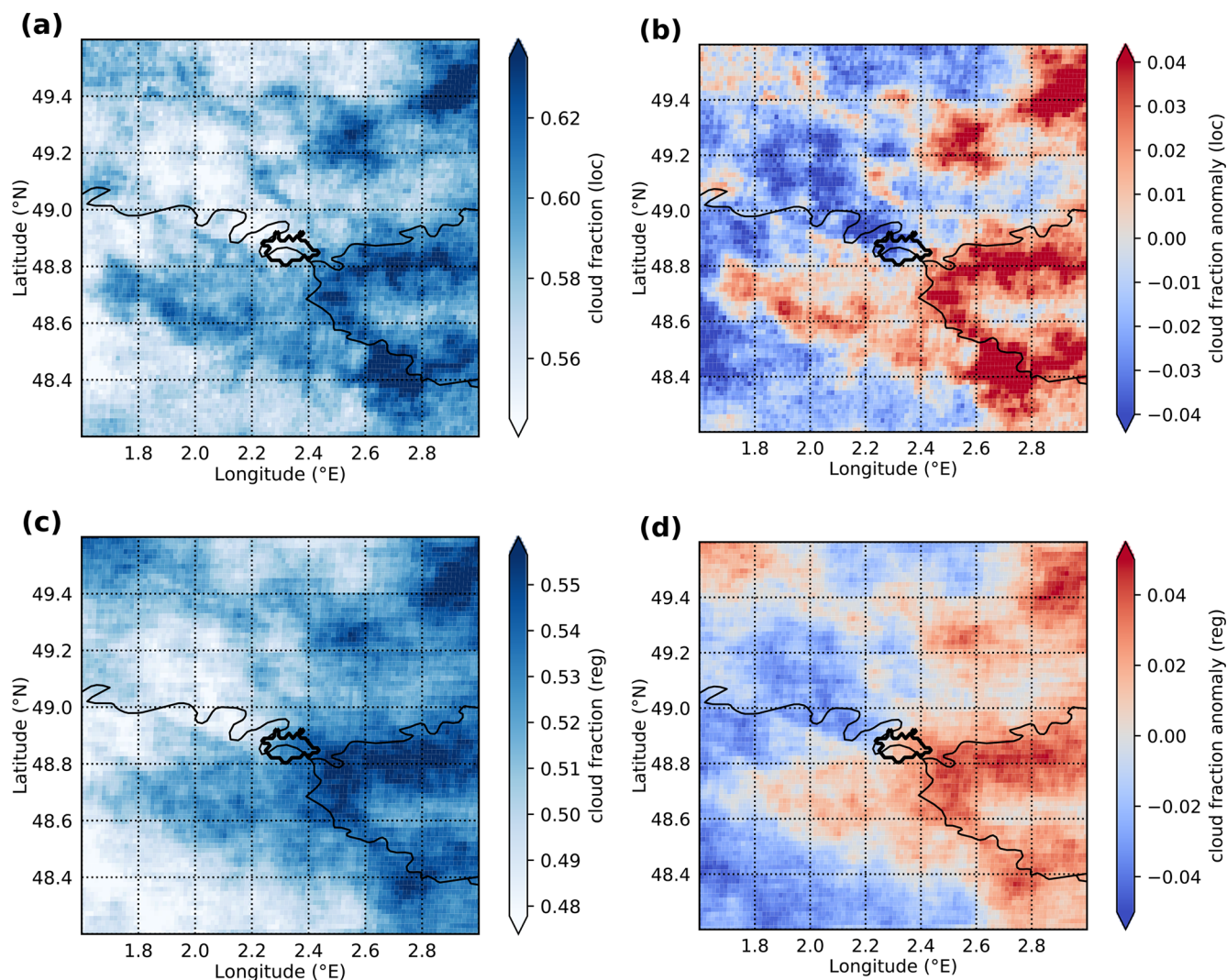
Figure 3. Cloud fraction CF_{loc} and CF_{reg} **(a, c)** and their spatial anomalies CF_{loc} anomaly and CF_{reg} anomaly **(b, d)** over the region of interest for the month of November during the years 2004–2019. Larger rivers are visualised as black lines, and the continuous urban area of Paris is shown as a black contour.

classes, Sect. 2.4) is expected to have a marginal influence on the cloud statistics.

4.2 Analysis of cloud patterns in typical fog conditions

Patterns of cloud clearings over urban areas are mostly expected in conditions of fog or low-stratus cloud (Gautam and Singh, 2018). In order to test the application of CM_{reg} to study spatial patterns of urban cloud modification, the derived cloud mask is filtered for specific boundary layer conditions using the ERA5 reanalysis data (cf. Sect. 2.3). Cloud fraction anomaly patterns constrained by these meteorological conditions show small-scale features that can be clearly associated with surface characteristics (Fig. 6). Notable is the distinct relative decrease of presumed fog and low-stratus cloud fraction directly over the centre of Paris and extend-

ing to its western edges on the order of $\sim 10\%$ relative to the regional average CF. This spatial pattern is usually associated with the dissipation of fog or low-stratus cloud by depletion of liquid water at the surface or the lifting of the cloud base (Wærsted et al., 2019; Williams et al., 2015; Underwood and Hansen, 2008; Haeffelin et al., 2005). This negative cloud fraction anomaly is likely a signal of the occurrences of fog holes (Gautam and Singh, 2018) over the urban centre of Paris. It is interesting to note that the negative anomaly extends to the west of the city, suggesting that in the high-pressure situations considered here winds may be predominating from easterly directions. The effect of wind direction on the cloud fraction anomaly is considered minor due the focus on conditions with low wind speeds only.

Positive cloud fraction anomalies show clear links to topography, especially over river valleys where fog is natu-
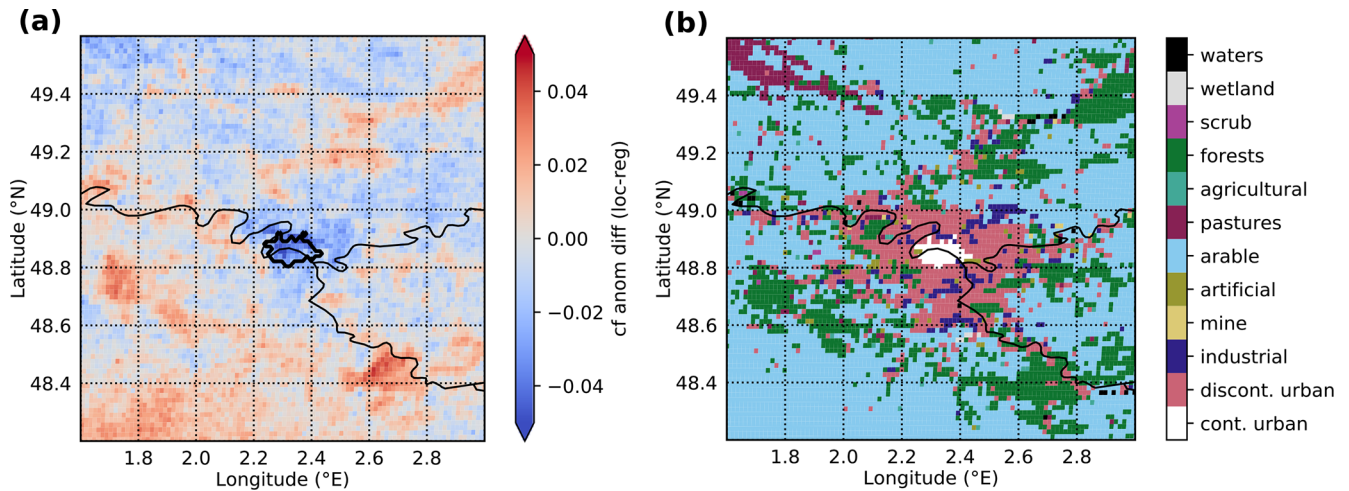
Figure 4. **(a)** Difference between cloud fraction anomalies (CF_{loc} anomaly $- CF_{reg}$ anomaly) from local and regional empirical cloud detection approaches. **(b)** Main CORINE land cover classes L2 and L3 for the year 2012. Larger rivers are visualised as black lines.
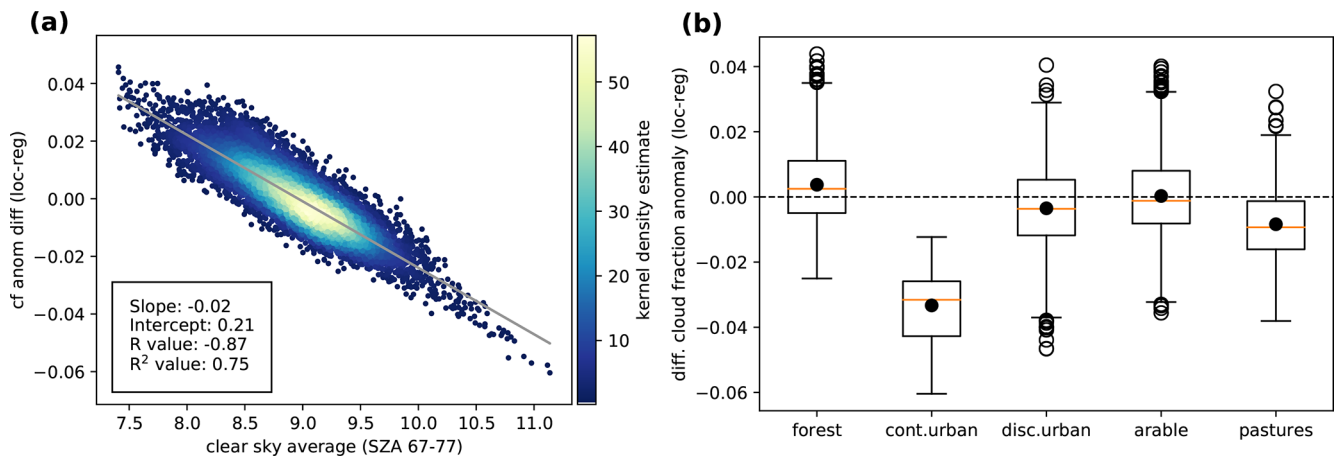

Figure 5. **(a)** Difference between cloud fraction anomalies (LECDA–RECDA) and clear-sky reflectance (average of SZA bins 67–77). Bright colours represent a higher probability density of the data points using Gaussian kernels. **(b)** Difference between cloud fraction anomalies (CF_{loc} anomaly $- CF_{reg}$ anomaly) from local and regional empirical cloud detection approaches grouped by five main CORINE land cover types: forest, continuous urban fabric, discontinuous urban fabric, arable land and pastures. The median and mean of each class are presented as horizontal orange lines and black dots, respectively. Whiskers are equal to 1.5 times the interquartile range beyond the first and third quartiles. Outliers are shown as black circles.

rally favoured (Bendix, 2002). This is visible in the northeast (Oise), the northwest (Epte, Avelon, Thérain), and east and southeast (Marne, Seine) of the study area. The high-resolution cloud fraction anomaly map thus highlights the ability of the regional approach to detect small-scale cloud patterns that can be linked to land cover characteristics and the topography.

For a systematic assessment of the land surface influences on spatial cloud fraction anomalies in the study region, cloud fraction anomalies are grouped by land cover types in Fig. 7. The land cover class "continuous urban" shows an average negative cloud fraction anomaly of 6 % that can be interpreted as a lower bound of the magnitude of the city's effect

on fog occurrence, as Paris is located in a river valley and should thus have favourable conditions for fog formation. Over pastures an increase by 4 % is notable; however, this is likely linked to the location of the pastures in the Avelon river valley in the NW of the study area. Smaller anomalies of the remaining land cover classes are partly explained by the combined influence of land cover type, terrain height and the dynamic of the fog hole evolution itself. Discerning the way that this cloud pattern is linked to the urban heat island of Paris, the land surface characteristics and any further meteorological variables will require further analysis beyond the scope of this paper.
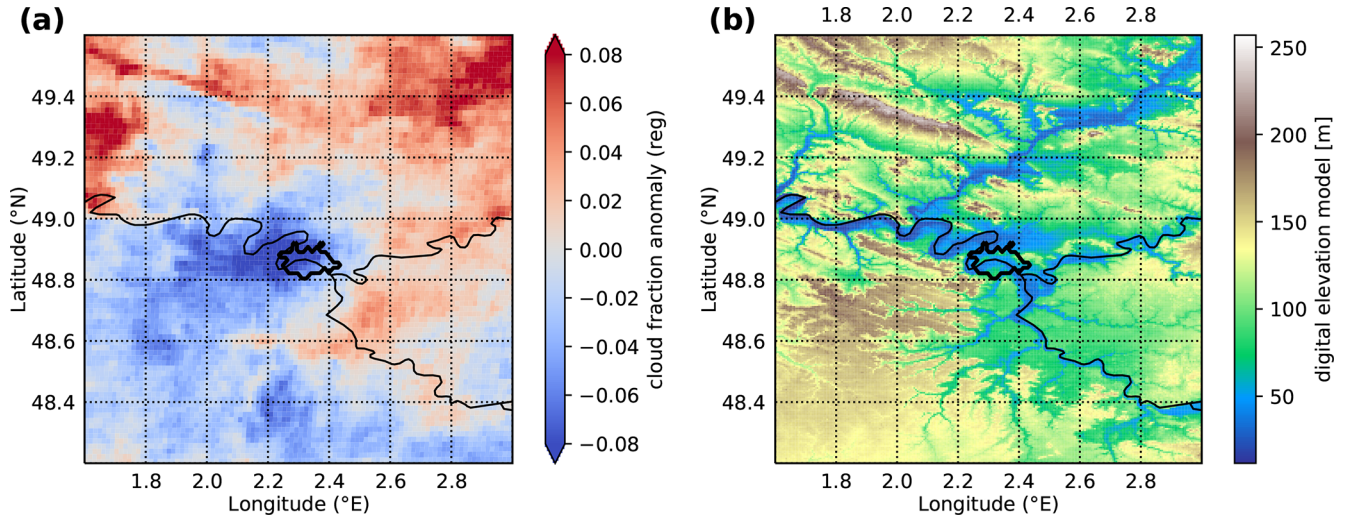
Figure 6. (a) Cloud fraction CF_{reg} anomaly (difference between the cloud fraction of the pixel and the average cloud fraction of the region) from the Regional Empirical Cloud Detection Approach constrained by the following meteorological conditions: low wind speed ($< 3\,\mathrm{m\,s^{-1}}$), low BLH ($< 300\,\mathrm{m}$) and high MSL ($> 1020\,\mathrm{hPa}$). **(b)** European Digital Elevation Model with a spatial resolution of 25 m. The Seine river is visualised as a black line.
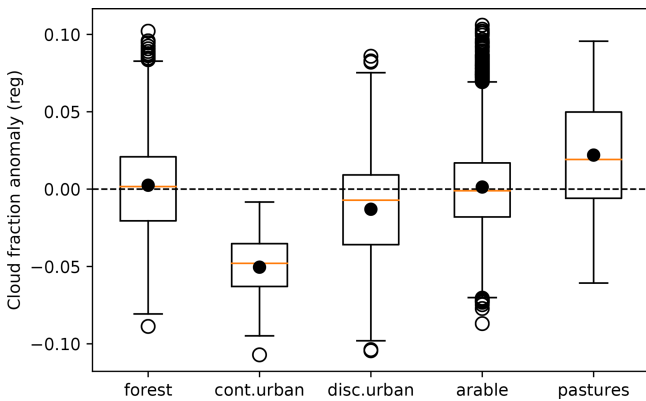

Figure 7. Cloud fraction CF_{reg} anomaly from the Regional Empirical Cloud Detection Approach constrained by meteorological conditions (Fig. 6) and grouped by five main CORINE land cover types: forest, continuous urban fabric, discontinuous urban fabric, arable land and pastures. Details are the same as in Fig. 5.

4.3 Validation with ground-based cloud fraction

The validation of the two cloud masks CM_{loc} and CM_{reg} with respect to the CloudNet data seeks to compare both approaches with ground truth data using standard measures of performance. However, it is not intended to prove the better applicability of one over another for the analysis of land surface effects on boundary layer clouds, as shown in Figs. 4a and 5a.

In Table 1 it is shown that both approaches perform similarly in terms of overall quality but achieve this with very different performance characteristics. In general, CM_{loc} outperforms CM_{reg} with a probability of detection (POD) of 0.77

Table 1. CloudNet validation results for the cloud masks using the local (CM_{loc}) and the regional (CM_{reg}) threshold. Validation measurements used are as follows: probability of detection (POD), false alarm ratio (FAR), percentage correct (PC), critical success index (CSI), bias score (BS) and Heidke skill score (HSS).

Class	POD	FAR	PC	CSI	BS	HSS
CM_{loc}	0.77	0.04	0.86	0.75	0.8	0.72
CM_{reg}	0.69	0.0	0.83	0.68	0.69	0.66

compared to 0.69. However, the false alarm ratio (FAR) of 0.0 for CM_{reg} falls below that for CM_{loc} of 0.04. These results are expected as T_{loc} is generally lower than T_{reg} (except where $T_{loc} = \max(T_{loc})$), leading to a general underestimation of clouds in CM_{reg} with the benefit of reducing false alarms. It is notable that a main difference in validating both approaches originates from the definition of a more cloud-conservative threshold vs. a more clear-conservative threshold.

The CloudNet data validate both approaches over a location with a surface reflectance of 8.65 (see Fig. 2) that is close to the scene average of 8.96. It is expected that the validation of the local and regional approach will give the same or nearly identical results over the bright surfaces of the city, while over darker forest regions the local approach is likely to have a much higher POD, with little influence on the FAR.

For assessing the general performance of both approaches (CM_{reg} and CM_{loc}) in the context of various thresholds, POD and FAR are computed along the range of HRV values and exemplary for different SZA (Fig. 8). The resulting pseudo-ROC (receiver operating characteristic) curves can be used
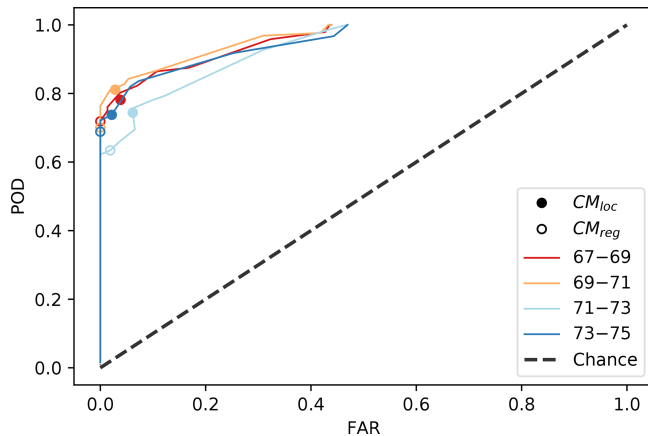
Figure 8. Relationship of POD and FAR for local and regional cloud masks (CM_{loc} and CM_{reg}; filled and transparent circles) and for a range of different HRV threshold values and four solar zenith angle (SZA) bins (coloured lines). HRV threshold values increase from the top to the bottom of the plot, i.e. from HRV minimum to HRV maximum per SZA. The pseudo-ROC curve shows the best performance of the HRV threshold values used for cloud masking in the upper-left corner, while accuracy decreases towards the dashed diagonal line.

to determine the ideal threshold for cloud detection, which typically can be found in the upper-left-hand corner with a minimal FAR and a maximal POD. The main outcome is that CM_{reg} for the SZA bins 67(–69) and 69(–71) is close to being an optimal tradeoff, achieving a low FAR while losing only little detection skill (POD) when compared to CM_{loc} and alternative HRV thresholds. The proposed regional approach, however, shows a lower POD compared to the local approach that can be attributed to the general cloud underestimation that is true for all the different surface reflectances equally. Further, for the application suggested in this study the ROC cannot provide a measure to evaluate the performance of LECDA compared to RECDA as done in Sect. 4.2.

The ground-based validation of the satellite-retrieved cloud masks is difficult as differences can be due to the specific measurement methods, including the different scales and the way cloud net data are aggregated to match the satellite data (cf. Greuell and Roebeling, 2009).

The performance of RECDA may vary depending on the location-specific characteristics of the land cover type, the satellite viewing geometry and the domain size. It is assumed that the performance of RECDA (similar to LECDA) will decrease in cases where the bimodal distribution broadens and flattens due to the effect of aerosols, thin liquid clouds, sub-pixel clouds or cloud edges contributing to the overlap area. It is expected that the presence of thin liquid clouds can lead to a surface-dependent bias in RECDA that is assumed to be minor compared to the influence of the surface signal in LECDA. Based on this study we can recommend RECDA for the proposed application and suggested domain size with

a clear-sky reflectance between 8 % and 10 % (see Fig. 2). RECDA will not provide reliable results over regions with clear-sky reflectances varying between 8 % and 50 %, e.g. agricultural and desert regions. In addition, the gradual approximation of cloud and clear-sky reflectance in the morning and evening twilight contributes to the flattening of the distribution curve and will degrade the ability of both algorithms to separate clouds and clear sky. The SZA dependence of the clear-sky reflection is known by other cloud mask algorithms as well, e.g. EUMETSAT (2019).

5 Conclusions

This study presents and compares the applicability of two empirical cloud-masking approaches based on the single HRV channel of MSG SEVIRI for the high-resolution analysis of land surface effects on boundary layer clouds. The performance of the cloud mask based on the regional approach RECDA is compared to the local approach LECDA with respect to the influence of the underlying surface. Both cloud mask approaches and obtained cloud fractions are analysed over a $\sim 150\,km \times 150\,km$ region centred on Paris, where holes in fog and low-stratus cloud are expected.

It is shown that cloud masks obtained from LECDA result in a relative underestimation of cloud occurrence over the bright urban surface of Paris (up to 5 %), while cloud occurrence is overestimated in relative terms over dark surfaces, e.g. forests (up to 5 %), when compared to RECDA. This leads to the conclusion that studies using such locally optimised approaches have to be attentive when associating different land surface classes to cloud occurrence.

The application of the regional approach RECDA to compute cloud masks in contrast to a pixel-based surface-dependent threshold is shown to be advantageous as local biases due to differences of surface reflectance can be mostly prevented. As suggested in the study's context, the cloud mask obtained from empirically based regional thresholds is more robust towards variability of the surface reflectance and shows a reduced FAR. Based on this approach, a reduction of cloud cover filtered for typical fog low-stratus conditions of up to 10 % could have been observed for the month of November over the urban area of Paris and to its west. This spatial cloud pattern is associated with the urban surface that probably affects the boundary layer through gradual heat release by the urban surface. In addition to this prominent decrease, small-scale cloud enhancements attributed to river valleys, topography and forest regions prove the plausibility of the cloud mask. It should be noted that the cloud masks obtained with RECDA are likely underestimating cloud occurrence, and thus their application is limited to the quantification and analysis of regional cloud cover anomalies.

This study shows the great potential of an HRV-derived robust cloud mask that is designed to be more land surface independent than locally optimised approaches, and it is thus

ideal for the satellite-based quantification of small-scale interactions of the urban surface and boundary layer clouds in the range of ~ 1–$2\,\mathrm{km}$. The relatively simple and independent implementation and application of this approach for regional and urban analyses is expected to have a diverse potential on an Europe-wide scale and needs to be tested for a varying domain size and temporally varying surface reflectance. Future research may also include the expansion and validation of this approach to regions that have experienced human-made or naturally induced land cover changes. Understanding the impact of vegetation changes on clouds, e.g. due to heat stress, is a relevant aspect in a changing climate and can be addressed by combining this cloud-masking approach with future satellite missions such as FLEX (FLuorescence EXplorer).

Appendix A

Table A1. CloudNet target classification

Class	Definition
0	Clear sky
1	Cloud liquid droplets only
2	Drizzle or rain
3	Drizzle or rain coexisting with cloud liquid droplets
4	Ice particles
5	Ice coexisting with supercooled liquid droplets
6	Melting ice particles
7	Melting ice particles coexisting with cloud liquid droplets
8	Aerosol particles, no cloud or precipitation
9	Insects, no cloud or precipitation
10	Aerosol coexisting with insects, no cloud or precipitation

Review statement. This paper was edited by Andrew Sayer and reviewed by Sebastian Bley and one anonymous referee.

References

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, J. Geophys. Res.-Atmos., 103, 32141–32157, https://doi.org/10.1029/1998JD200032, 1998.

Andersen, H. and Cermak, J.: First fully diurnal fog and low cloud satellite detection reveals life cycle in the Namib, Atmos. Meas. Tech., 11, 5461–5470, https://doi.org/10.5194/amt-11-5461-2018, 2018.

Bendix, J.: A satellite-based climatology of fog and low-level stratus in Germany and adjacent areas, Atmos. Res., 64, 3–18, https://doi.org/10.1016/S0169-8095(02)00075-3, 2002.

Bley, S. and Deneke, H.: A threshold-based cloud mask for the high-resolution visible channel of Meteosat Second Generation SEVIRI, Atmos. Meas. Tech., 6, 2713–2723, https://doi.org/10.5194/amt-6-2713-2013, 2013.

Cermak, J.: SOFOS – A New Satellite-based Operational Fog Observation Scheme, PhD thesis, Philipps-Universität Marburg, http://archiv.ub.uni-marburg.de/diss/z2006/0149/ (last access: 1 October 2021), 2006.

Cermak, J. and Bendix, J.: A novel approach to fog/low stratus detection using Meteosat 8 data, Atmos. Res., 87, 279–292, https://doi.org/10.1016/j.atmosres.2007.11.009, 2008.

Chen, P. Y., Srinivasan, R., and Fedosejevs, G.: An automated cloud detection method for daily NOAA 16 advanced very high resolution radiometer data over Texas and Mexico, J. Geophys. Res.-Atmos., 108, 1–9, https://doi.org/10.1029/2003jd003554, 2003.

Collier, C. G.: The impact of urban areas on weather, Q. J. Roy. Meteor. Soc., 132, 1–25, https://doi.org/10.1256/qj.05.199, 2006.

Deneke, H., Feijt, A., van Lammeren, A., and Simmer, C.: Validation of a physical retrieval scheme of solar surface irradiances from narrowband satellite radiances, J. Appl. Meteorol., 44, 1453–1466, https://doi.org/10.1175/JAM2290.1, 2005.

Deneke, H. M. and Roebeling, R. A.: Downscaling of METEOSAT SEVIRI 0.6 and 0.8 μm channel radiances utilizing the high-resolution visible channel, Atmos. Chem. Phys., 10, 9761–9772, https://doi.org/10.5194/acp-10-9761-2010, 2010.

Deneke, H. M., Knap, W. H., and Simmer, C.: Multiresolution analysis of the temporal variance and correlation of transmittance and reflectance of an atmospheric column, J. Geophys. Res.-Atmos., 114, D17206, https://doi.org/10.1029/2008JD011680, 2009.

Derrien, M., Gléau, H. L., and Raoul, M.: The use of the high resolution visible in SAFNWC/MSG cloud mask, Proceedings of the 2010 EUMETSAT Meteorological Satellite Conference, 20–24 September 2010, Córdoba, Spain, 57 pp., http://hal.archives-ouvertes.fr/meteo-00604325/ (last access: 5 July 2021), 2010.

Di Vittorio, A. V. and Emery, W. J.: An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land, IEEE T. Geosci. Remote, 40, 1682–1694, https://doi.org/10.1109/TGRS.2002.802455, 2002.

Dozier, J. and Painter, T. H.: Multispectral and hyperspectral remote sensing of alpine snow properties, Annu. Rev. Earth Pl. Sc., 32, 465–494, https://doi.org/10.1146/annurev.earth.32.101802.120404, 2004.

EEA: Copernicus Land Monitoring Service – Reference Data: EU-DEM, https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1 (last access: 25 November 2021), 2017.

EEA: CORINE Land Cover – User Manual, Copernicus Land Monitoring Service, https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012?tab=mapview, last access: 25 November 2021.

EUMETSAT: EUMETSAT CM SAF Algorithm Theoretical Basis Document (ATBD) Meteosat Solar Surface Radiation and effective Cloud Albedo Climate Data Records - Heliosat (SARAH-2), Satellite Application Facility on Climate Monitoring (CM SAF), 31 pp., https://doi.org/10.5676/EUM_SAF_CM/SARAH/V002_01, 2019 (data available at: https://www.eumetsat.int/access-our-data, last access: 22 April 2020).

Fan, J., Wang, Y., Rosenfeld, D., Liu, X., Fan, J., Wang, Y., Rosenfeld, D., and Liu, X.: Review of Aerosol-Cloud Interactions: Mechanisms, Significance, and Challenges, J. Atmos. Sci., 73, 4221–4252, https://doi.org/10.1175/JAS-D-16-0037.1, 2016.

Gautam, R. and Singh, M. K.: Urban Heat Island Over Delhi Punches Holes in Widespread Fog in the Indo-Gangetic Plains, Geophys. Res. Lett., 45, 1114–1121, https://doi.org/10.1002/2017GL076794, 2018.

Greuell, W. and Roebeling, R. A.: Toward a standard procedure for validation of satellite-derived cloud liquid water path: A study with SEVIRI data, J. Appl. Meteorol. Clim., 48, 1575–1590, https://doi.org/10.1175/2009JAMC2112.1, 2009.

Haeffelin, M., Barthès, L., Bock, O., Boitel, C., Bony, S., Bouniol, D., Chepfer, H., Chiriaco, M., Cuesta, J., Delanoë, J., Drobinski, P., Dufresne, J.-L., Flamant, C., Grall, M., Hodzic, A., Hourdin, F., Lapouge, F., Lemaître, Y., Mathieu, A., Morille, Y., Naud, C., Noël, V., O'Hirok, W., Pelon, J., Pietras, C., Protat, A., Romand, B., Scialom, G., and Vautard, R.: SIRTA, a ground-based atmospheric observatory for cloud and aerosol research, Ann. Geophys., 23, 253–275, https://doi.org/10.5194/angeo-23-253-2005, 2005.

Hall, D. K., Riggs, G. A., and Salomonson, V. V.: Algorithm theoretical basis document (ATBD) for the MODIS snow and sea ice-mapping algorithms, Tech. rep., NASA (National Aeronautics and Space Administration), Goddard Space Flight Center, Greenbelt, MD, 2001.

Han, J.-Y., Baik, J.-J., and Lee, H.: Urban impacts on precipitation, Asia-Pac. J. Atmos. Sci., 50, 17–30, https://doi.org/10.1007/s13143-014-0016-7, 2014.

Henken, C. C., Schmeits, M. J., Deneke, H., and Roebeling, R. A.: Using MSG-SEVIRI cloud physical properties and weather radar observations for the detection of Cb/TCu clouds, J. Appl. Meteorol. Clim., 50, 1587–1600, https://doi.org/10.1175/2011JAMC2601.1, 2011.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.adbb2d47, 2018 (data available at: https://cds.climate.copernicus.eu/cdsapp#!/

dataset/reanalysis-era5-single-levels?tab=overview, last access: 2 November 2021).

Hutchison, K. D., Roskovensky, J. K., Jackson, J. M., Heidinger, A. K., Kopp, T. J., Pavolonis, M. J., and Frey, R.: Automated cloud detection and classification of data collected by the Visible Infrared Imager Radiometer Suite (VIIRS), Int. J. Remote Sens., 26, 4681–4706, https://doi.org/10.1080/01431160500196786, 2005.

Illingworth, A. J., Hogan, R. J., O'Connor, E. J., Bouniol, D., Brooks, M. E., Delanoë, J., Donovan, D. P., Eastment, J. D., Gaussiat, N., Goddard, J. W., Haeffelin, M., Klein Baltink, H., Krasnov, O. A., Pelon, J., Piriou, J. M., Protat, A., Russchenberg, H. W., Seifert, A., Tompkins, A. M., van Zadelhoff, G. J., Vinit, F., Willen, U., Wilson, D. R., and Wrench, C. L.: Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, B. Am. Meteorol. Soc., 88, 883–898, https://doi.org/10.1175/BAMS-88-6-883, 2007 (data available at: https://actris.nilu.no/, last access: 2 February 2020).

Ipe, A., Clerbaux, N., Bertrand, C., Dewitte, S., and Gonzalez, L.: Pixel-scale composite top-of-the-atmosphere clear-sky reflectances for Meteosat-7 visible data, J. Geophys. Res.-Atmos., 108, 1–7, https://doi.org/10.1029/2002jd002771, 2003.

Jain, A., Duin, R. P. W., and Mao, J.: Statistical pattern recognition: a review, IEEE T. Pattern Anal., 22, 4–37, https://doi.org/10.1109/34.824819, 2000.

Kim, M., Cermak, J., Andersen, H., Fuchs, J., and Stirnberg, R.: A new satellite-based retrieval of low-cloud liquid-water path using machine learning and meteosat seviri data, Remote Sens.-Basel, 12, 1–16, https://doi.org/10.3390/rs12213475, 2020.

Klüser, L., Rosenfeld, D., Macke, A., and Holzer-Popp, T.: Observations of shallow convective clouds generated by solar heating of dark smoke plumes, Atmos. Chem. Phys., 8, 2833–2840, https://doi.org/10.5194/acp-8-2833-2008, 2008.

Liang, X., Miao, S., Li, J., Bornstein, R., Zhang, X., Gao, Y., Chen, F., Cao, X., Cheng, Z., Clements, C., Dabberdt, W., Ding, A., Ding, D., Dou, J. J., Dou, J. X., Dou, Y., Grimmond, C. S., Gonzál, J. E., He, J., Huang, M., Huang, X., Ju, S., Li, Q., NiYogi, D., Quan, J., Sun, J., Sun, J. Z., Yu, M., Zhang, J., Zhang, Y., Zhao, X., Zheng, Z., and Zhou, M.: Surf: Understanding and predicting urban convection and haze, B. Am. Meteorol. Soc., 99, 1391–1413, https://doi.org/10.1175/BAMS-D-16-0178.1, 2018.

Matthews, E. and Rossow, W. B.: Regional and Seasonal Variations of Surface Reflectance from Satellite Observations at 0.6 µm, J. Clim. Appl. Meteorol., 26, 170–202, https://doi.org/10.1175/1520-0450(1987)026<0170:RASVOS>2.0.CO;2, 1987.

Minnis, P. and Harrison, E. F.: Diurnal variability of regional cloud and clear-sky radiative parameters derived from GOES data. Part III: November 1978 radiative parameters., J. Clim. Appl. Meteorol., 23, 1032–1051, https://doi.org/10.1175/1520-0450(1984)023<1032:DVORCA>2.0.CO;2, 1984.

Nilo, S. T., Romano, F., Cermak, J., Cimini, D., Ricciardelli, E., Cersosimo, A., Di Paola, F., Gallucci, D., Gentile, S., Geraldi, E., Larosa, S., Ripepi, E., and Viggiano, M.: Fog detection based on Meteosat Second Generation-Spinning enhanced visible and infrared imager high resolution visible channel, Remote Sens.-Basel, 10, 1–20, https://doi.org/10.3390/rs10040541, 2018.

Pauli, E., Cermak, J., and Teuling, A. J.: Enhanced nighttime fog and low stratus occurrence over the Landes forest, France, Geophys. Res. Lett., 49, e2021GL097058, https://doi.org/10.1029/2021GL097058, 2022.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, https://doi.org/10.1007/s13398-014-0173-7.2, 2011.

Richter, R., Louis, J., and Berthelot, B.: Sentinel-2 MSI – Level 2A Products Algorithm Theoretical Basis Document, European Space Agency, ESA SP Publ., 49, 1–72, 2012.

Rossow, W. B. and Garder, L. C.: Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP, J. Climate, 6, 2341–2369, https://doi.org/10.1175/1520-0442(1993)006<2341:CDUSMO>2.0.CO;2, 1993.

Saunders, R. W. and Kriebel, K. T.: An improved method for detecting clear sky and cloudy radiances from AVHRR data, Int. J. Remote Sens., 9, 123–150, https://doi.org/10.1080/01431168808954841, 1988.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: AN INTRODUCTION TO METEOSAT SECOND GENERATION (MSG), B. Am. Meteorol. Soc., 83, 977–992, https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2, 2002.

Schulz, H. M., Thies, B., Cermak, J., and Bendix, J.: 1 km fog and low stratus detection using pan-sharpened MSG SEVIRI data, Atmos. Meas. Tech., 5, 2469–2480, https://doi.org/10.5194/amt-5-2469-2012, 2012.

Schutgens, N. A. and Roebeling, R. A.: Validating the validation: the influence of liquid water distribution in clouds on the intercomparison of satellite and surface observations, J. Atmos. Ocean. Tech., 26, 1457–1474, https://doi.org/10.1175/2009JTECHA1226.1, 2009.

Shepherd, J. M.: A Review of Current Investigations of Urban-Induced Rainfall and Recommendations for the Future, Earth Interact., 9, 1–27, https://doi.org/10.1175/EI156.1, 2005.

Stowe, L. L., Davis, P. A., and Mcclain, E. P.: Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer, J. Atmos. Ocean. Tech., 16, 656–681, https://doi.org/10.1175/1520-0426(1999)016<0656:SBAIEO>2.0.CO;2, 1999.

Strabala, K. I., Ackerman, S. A., and Menzel, W. P.: Cloud Properties inferred from 8–12-µm Data, J. Appl. Meteorol., 33, 212–229, https://doi.org/10.1175/1520-0450(1994)033<0212:CPIFD>2.0.CO;2, 1994.

Teuling, A. J., Taylor, C. M., Meirink, J. F., Melsen, L. A., Miralles, D. G., van Heerwaarden, C. C., Vautard, R., Stegehuis, A. I., Nabuurs, G.-J., and de Arellano, J. V.-G.: Observational evidence for cloud cover enhancement over western European forests, Nat. Commun., 8, 14065, https://doi.org/10.1038/ncomms14065, 2017.

Theeuwes, N. E., Barlow, J. F., Teuling, A. J., Grimmond, C. S. B., and Kotthaus, S.: Persistent cloud cover over mega-cities linked to surface heat release, npj Climate and Atmospheric Science, 2, 15, https://doi.org/10.1038/s41612-019-0072-x, 2019.

Tøttrup, C.: EU-DEM Statistical Validation Report August 2014, European Environment Agency (EEA), Copenhagen, Denmark, p. 27, 2014.

Underwood, S. J. and Hansen, C.: Investigating urban clear islands in fog and low stratus clouds in the San Joaquin Valley of California, Phys. Geogr., 29, 442–454, https://doi.org/10.2747/0272-3646.29.5.442, 2008.

Varentsov, M., Wouters, H., Platonov, V., and Konstantinov, P.: Megacity-Induced Mesoclimatic Effects in the Lower Atmosphere: A Modeling Study for Multiple Summers over Moscow, Russia, Atmosphere, 9, 50, https://doi.org/10.3390/atmos9020050, 2018.

Veefkind, J. P., de Haan, J. F., Sneep, M., and Levelt, P. F.: Improvements to the OMI O_2-O_2 operational cloud algorithm and comparisons with ground-based radar–lidar observations, Atmos. Meas. Tech., 9, 6035–6049, https://doi.org/10.5194/amt-9-6035-2016, 2016.

Wærsted, E. G., Haeffelin, M., Steeneveld, G., and Dupont, J.: Understanding the dissipation of continental fog by analysing the LWP budget using idealized LES and in situ observations, Q. J. Roy. Meteor. Soc., 145, 784–804, https://doi.org/10.1002/qj.3465, 2019.

Westerhuis, S., Fuhrer, O., Cermak, J., and Eugster, W.: Identifying the key challenges for fog and low stratus forecasting in complex terrain, Q. J. Roy. Meteor. Soc., 146, 3347–3367, https://doi.org/10.1002/qj.3849, 2020.

Williams, A. P., Schwartz, R. E., Iacobellis, S., Seager, R., Cook, B. I., Still, C. J., Husak, G., and Michaelsen, J.: Urbanization causes increased cloud base height and decreased fog in coastal Southern California, Geophys. Res. Lett., 42, 1527–1536, https://doi.org/10.1002/2015GL063266, 2015.

Yang, Y., Di Girolamo, L., and Mazzoni, D.: Selection of the automated thresholding algorithm for the Multi-angle Imaging SpectroRadiometer Radiometric Camera-by-Camera Cloud Mask over land, Remote Sen. Environ., 107, 159–171, https://doi.org/10.1016/j.rse.2006.05.020, 2007.

Zhong, S., Qian, Y., Zhao, C., Leung, R., Wang, H., Yang, B., Fan, J., Yan, H., Yang, X.-Q., and Liu, D.: Urbanization-induced urban heat island and aerosol effects on climate extremes in the Yangtze River Delta region of China, Atmos. Chem. Phys., 17, 5439–5457, https://doi.org/10.5194/acp-17-5439-2017, 2017.