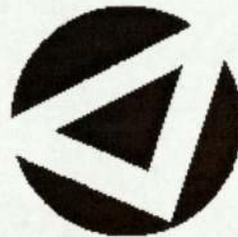


# Space and Time Modelling of Clouds.

THOMAS BERMUDEZ

MSc by Research in Pattern Analysis and Neural Networks



ASTON UNIVERSITY

August 2004

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

# Acknowledgements

I would like to thank my supervisors, Dan Cornford and Ian Nabney, for their availability, their help, their patience and especially Dan for his precious advice and encouragement. Working with you has been a real pleasure.

I would also like to thank all the people from the Neural Computing Research Group department for the quality of their lectures, their support and their friendly presence. Many thanks to Miss Vicky Bond. Thanks to my Master's fellow students: Christophe, Youssef, Kiriko, Dharmesh, Youenn and Pierre. Spending a year in such an environment has been really nice despite hard times.

Last but not least, I would like to thank my family for their love and support. *Incarnación et Antoine, Violette et Raymond, merci pour votre précieux soutien et vos encouragements. Violette et Raymond, j'apprécie à sa juste valeur l'effort que vous avez fait en m'accompagnant.* Thanks to my parents Joëlle et José, I owe you so much.

ASTON UNIVERSITY

# Space and Time Modelling of Clouds.

THOMAS BERMUDEZ

MSc by Research in Pattern Analysis and Neural Networks, 2004

## Thesis Summary

Satellite images provide detailed spatial and temporal information about the structure of the atmosphere: state at a time  $t$  or evolution from time  $t - n$  to time  $t$ . The conception of a model able to forecast the motion of the clouds (in other terms, a space-time model for clouds) relies on high-frequency space and time measurements of relevant features of the atmosphere. In the framework of this thesis, position, shape and brightness of the clouds are extracted from METEOSAT-6's raw visible images via a preprocessing step consisting in the setting of a static threshold on the 'cloud optical thickness' of each pixel in the image.

The next step in the project is the creation of a model for clouds. As clouds are assumed to be groups of 'cloud cells', modelling a cloud comes down to modelling each of its basis 'cloud cells'. Thus, a cloud is modelled as a Radial Basis Function network with Gaussian radial basis functions.

In the last step of the project, the dynamics of the model are investigated. The advection field, responsible (among other processes) for the moving of the cloud field, can be seen as a 'wind flow'. Forecasting the state of the motion of the clouds from time  $t - 1$  to time  $t$  consists in forecasting the next state of the advection field as well as estimating the parameters of the Gaussian radial basis functions modelling the 'cloud cells'. To do so, a Kalman filter-like approach has been undertaken: the filter supports estimations of the past, the present and even future states without knowing the precise nature of the modelled system.

**Keywords:** cloud winds, remote sensing, cloud modelling, space-time model

# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	Image processing . . . . .	6
1.2	Data assimilation . . . . .	8
1.3	Thesis overview . . . . .	9
<b>2</b>	<b>Description of the dataset</b>	<b>10</b>
2.1	Acquisition process . . . . .	10
2.2	Dataset . . . . .	12
2.3	Summary . . . . .	12
<b>3</b>	<b>Image preprocessing</b>	<b>13</b>
3.1	Observation model . . . . .	13
3.2	Computation of the solar flux density . . . . .	14
3.3	Digital to geographic coordinates . . . . .	14
3.3.1	Digital coordinate geometry . . . . .	14
3.3.2	Digital to geographic conversion routine . . . . .	15
3.3.3	Test of the conversion routine . . . . .	15
3.4	Computation of the reference brightness . . . . .	15
3.4.1	Gaussian Mixture Model . . . . .	17
3.4.2	Minimum brightness . . . . .	19
3.5	Discrimination cloudy/non-cloudy regions . . . . .	20
3.6	Discussion . . . . .	23
3.7	Summary . . . . .	24
<b>4</b>	<b>Modelling of the clouds</b>	<b>25</b>
4.1	Radial Basis Function network: framework . . . . .	25
4.2	Network initialisation . . . . .	26
4.3	Results . . . . .	27
4.4	Summary . . . . .	32
<b>5</b>	<b>Dynamics of the model</b>	<b>33</b>
5.1	Advection equation . . . . .	33
5.2	State-space model: Kalman filter . . . . .	34
5.2.1	Evolution of $\gamma$ . . . . .	35
5.2.2	Evolution of $\mathbf{v}$ . . . . .	35
5.2.3	Assimilation of $\gamma$ . . . . .	35
5.2.4	Assimilation of $\mathbf{v}$ . . . . .	37
5.3	Running the space and time model of clouds . . . . .	37
5.3.1	Initialisation of the model . . . . .	37
5.3.2	Forecast . . . . .	38
5.3.3	Results . . . . .	39
5.4	Discussion . . . . .	45
5.5	Summary . . . . .	46

CONTENTS

<b>6 Conclusion</b>	<b>47</b>
6.1 Summary of the work done . . . . .	47
6.2 Results . . . . .	47
6.3 Further work . . . . .	48
6.4 Afterword . . . . .	49
<b>A Digital to geographic coordinates conversion routine</b>	<b>50</b>
<b>B Maximum Likelihood and EM algorithm</b>	<b>53</b>
<b>C Initialisation of the network: example on a single window</b>	<b>55</b>
<b>D Kalman filter</b>	<b>59</b>
D.1 Time update: Evolution step . . . . .	59
D.2 Measurement update: Assimilation step . . . . .	60

# List of Figures

2.1	European sector. . . . .	11
2.2	European sector: visible satellite image. . . . .	11
3.1	Layout of the coastline of the European region. . . . .	16
3.2	Plotting of the geodetic coordinates (black points) of the coastline on a projection of the Earth viewed by a geostationary satellite (detail). . . . .	16
3.3	Histograms of brightness and first component of a three component Gaussian Mixture Model for two pixels. . . . .	18
3.4	Standard reflectance of the ground of the European sector (three component Gaussian mixture model case). . . . .	19
3.5	Minimum brightness of the ground of the European sector. . . . .	20
3.6	Example of discrimination cloudy/non-cloudy regions for the 13 <sup>th</sup> of March 2004 at 12.00 (GMT). . . . .	21
3.7	Example of discrimination cloudy/non-cloudy regions for the 14 <sup>th</sup> of March 2004 at 10.00 (GMT). . . . .	22
4.1	RBF network used for the thesis: a cloud is modelled as a mixture of Gaussian radial basis functions. . . . .	26
4.2	European sub-sector: $\gamma$ values. . . . .	27
4.3	Splitting of the original image in windows of size 50x50 pixels. . . . .	28
4.4	Split image and overlap. . . . .	29
4.5	Weight function. . . . .	29
4.6	Modelled cloud after reunification of the optimised ‘windows’. . . . .	30
4.7	Modelled cloud after reunification of the optimised ‘windows’: crosses correspond to the centres of the modelled ‘cloud cells’, different colours correspond to different ‘windows’. . . . .	31
5.1	The <i>time update</i> projects the current state estimate ahead in time. The <i>measurement update</i> adjusts the projected estimate by an actual measurement at that time. . . . .	35
5.2	Actual state of the process at time $t$ and initialisation of the model (first step). . . . .	40
5.3	Actual state of the process at time $t + 1$ and initialisation of the model (second step). . . . .	41
5.4	Forecast of the process at time $t + 3$ and actual state of the process at time $t + 3$ . . . . .	42
5.5	Forecast of the process at time $t + 4$ and actual state of the process at time $t + 4$ . . . . .	43
5.6	Forecast of the process at time $t + 5$ and actual state of the process at time $t + 5$ . . . . .	44
A.1	Viewing line from the spacecraft: one intersection. O is the centre of the Earth, S the satellite, I the observed point and P the projection of I in the Cartesian co-ordinate system of the Earth. . . . .	50
C.1	Original weighted ‘window’. . . . .	55
C.2	Removing of the effect of the first radial basis function. . . . .	56
C.3	Removing of the effect of the two first radial basis functions. . . . .	56
C.4	Removing of the effect of the three first radial basis functions. . . . .	57
C.5	Stopping criterion reached. . . . .	57
C.6	Moving of the centres after the global optimisation of the network has been performed. . . . .	58
C.7	Modelled cloud and centres of the modelled ‘cloud cells’. . . . .	58

# Chapter 1

## Introduction

In 2002, Cornford [5] and Batail [3] developed a statistical model for short-term rainfall forecasting on a Bayesian framework similar to the Kalman filter. Barillec [2] improved the model (by speeding it up using a variational Bayes method) in 2003. A first attempt to apply the model to the ‘cloud field’ (that is to say to create a space and time model for clouds) has been carried out the same year by Lê [8]. This thesis extends the earlier work, incorporating an improved model for the observation process and incorporates a procedure for parallelising the computation.

A short overview of the short term forecasting of the motion of the clouds as well as its constraints is given in the next paragraphs.

### 1.1 Image processing

The short-term forecasting of the motion of the clouds relies on an important but complex source of information: satellite images. The latter have to be processed in order to extract relevant features.

In the attempt to determine cloud spatial and temporal characteristics from satellite imagery for use in a neural network cloud classification scheme, Lewis *et al.* [9] used a Multilayer Perceptron neural network to segment the images (*cloud object definition* step). The adaptive threshold method used was based on a neural network analysis of the brightness histograms of sub-scenes of satellite images. The brightness histogram of each sub-scene being multi-modal in shape, the modes were assumed to consist of pixels characteristic of particular regions in the image: the first mode representing the sea-surface, the final mode representing bright clouds and modes between these representing the land-surface and clouds of varying intensity. The adaptive threshold algorithm was implemented using a Multilayer Perceptron (MLP) neural network with a single hidden layer. The back-propagation learning algorithm with an added term was used to train the MLP by minimising the network error energy function. The network architecture consisted of sixteen nodes (one for each sub-scene in the image) in the input layer, ten nodes in the hidden layer and one node in the output layer. The test set was then used to determine the performance of the network, which provided satisfactory results.

Simpson [14] made an inventory of the difficulties encountered with the accurate cloud detection in satellite scenes. The simplest approach for cloud detection in a scene is to apply a set of static thresholds (albedo, temperature. . .) for every pixel in the scene, nevertheless this method can fail for several reasons: sub-pixel clouds and cloud-pixel misalignment can lead to errors as can variations in brightness temperature, sensor aging and the variation of spectral response of clouds with type and height. Spatial coherence methods present an advantage over static threshold methods because they use the local spatial structure of the infrared radiance field to determine cloud-free and cloud-covered pixels. However, those methods can also fail because of the multilayered nature of clouds or the size of the clouds in the scene being smaller than the instrument's field of view. A hybrid cloud detection procedure has been developed following the former observations (next paragraph).

Simpson *et al.* [15] developed a new algorithm for accurate cloud detection in Advanced Very High Resolution Radiometer (AVHRR) scenes over land. The AVHRR Split-and-Merge Clustering (ASMC) algorithm for cloud detection in AVHRR scenes over land consists of two steps:

1. a split-and-merge clustering of the input data which segments the scene into its natural groupings;
2. a cluster-labelling procedure that uses scene-specific, joint three-dimensional adaptive labelling thresholds to label the clusters as either cloud, cloud-free land, or uncertain.

Results obtained with the ASMC algorithm compare favourably with those obtained from a wide range of currently used algorithms to detect cloud over land in AVHRR data. Nevertheless, the algorithm cannot perform the segmentation of ice and / or snow from clouds.

Mukherjee *et al.* [11] deal with the problem of cloud tracking within a sequence of geo-stationary satellite images. The tracking of clouds involves a major problem: a homogeneous cloud mass needs to be detected for tracking. The detection of the cloud mass boundary is a problem of image segmentation and a scale-sensitive image classification approach has been introduced: the classification is performed on a scale space representation of the satellite imagery (set of scaled images). The classifier detects homogeneous cloud segments with minimum intra-segment classification error. The scale space approach provided good results.

Papin *et al.* [13] investigated the early and accurate segmentation of low clouds during the nighttime. The information supplied by the single infrared METEOSAT channel available at night not being sufficient to discriminate between low clouds and ground during night from a single image, several sources of information have been extracted from an infrared image sequence. Both relevant local motion-based measurements, intensity images and thermal parameters estimated over blocks, along with local contextual information have been exploited. A statistical contextual labelling process in two classes "low clouds" and "clear sky" is performed on the warmer pixels. It is formulated within a Bayesian estimation framework associated with Markov random field (MRF) models which comes down to minimising a global energy function composed of three terms: two data-driven terms (thermal

and motion-based ones) and a regularization term expressing *a priori* knowledge on the label field. Experimental results on representative meteorological situations favourably compare with AVHRR cloud classifications.

In the framework of the thesis, the approach taken by Lewis *et al.* [9] has been followed: given a sequence of images, histograms of brightness of the scene have been plotted leading to the discovery of characteristic features. The latter, corresponding to modes in the histograms, have been identified with the distribution of brightness of different classes of elements. Then a static threshold has been set up to discriminate the cloudy parts of the satellite images from the non-cloudy ones. This technique has proved its efficiency in the project but also its limits: applying a threshold to a scene as important as the European sector can lead to a poor discrimination at its edges as will be seen in Chapter 3.

## 1.2 Data assimilation

In the physical sciences such as meteorology or oceanography, combining measurements with the dynamics of the underlying models is usually referred to as *data assimilation*. Data assimilation improves the reconstruction of the image fields of interest.

Asif [1] motivated by the reconstruction of image fields like ocean velocity or sea surface height in physical oceanography, developed algorithms to couple the dynamics of the fields (often modelled by partial differential equations) with sparse measurements. The main approaches to the space-time dependent physical systems, which Asif was interested in, are of two classes : optimal interpolation methods and least square methods. The least square methods minimize in a statistical sense the data misfit under the constraint of the model dynamics. These include the Kalman-Bucy filter (KBf, more often called Kalman filter). Assimilating data with this algorithm is challenging due to its computational cost. Asif developed four efficient implementations of the KBf that reduce its computational cost by combining the block structure of the underlying dynamical models and the sparseness of the measurements.

Le Marshall *et al.* [10] have addressed the tropical cyclone track forecasting problem in three ways. First, modelling and data assimilation have been performed at very high resolution. Second, a newly developed data source, high spatial and temporal resolution cloud drift winds, has been used to augment the poor data base upon which forecasts are based. Finally, a range of continuous assimilation techniques have been tested. Le Marshall *et al.* concluded that a continuous assimilation, combined with substantial high resolution data base and high resolution modelling have the potential to greatly improve the accuracy of tropical cyclone track forecasts.

Ohring *et al.* [12] describe the National Centers for Environmental Prediction (NCEP) global and regional data assimilation system. The latter uses a variational analysis technique which allows the incorporation of diverse types of observations, model forecasts fields and physical and dynamical constraints into a single analysis while accounting for the statistical properties of the data and model

forecast. The model forecast is used to extrapolate the information in observations at previous times to the current analysis time. Like all variational schemes, the NCEP analysis system minimizes an objective function comprising terms of analysis, of background model and of observations. One of the major problems identified by Ohring *et al.* lies in the specification of the forecast model's background errors and observation errors. The model background generally consists of a short forecast that is used at the first guess for the analysis field. Error information is needed to allow proper weighting of the background field and observations.

Lewis *et al.* [9] used a method developed on the Hopfield neural network to match cloud objects across sequential images. The Hopfield network is composed of neurons linked to other neurons and biases by weighted connections. However, its main characteristic comes from the fact that it is recurrent: neuron outputs are fed back to neuron inputs. Once matched, temporal parameters were extracted from corresponding clouds. Parameters describing the change of shape of an object with respect to time and parameters describing the visible spectral signature of cloud samples were derived from the shape parameters for corresponding clouds. The cloud-wind vectors generated from the results of the Hopfield network matching technique were shown to be good estimates of the direction of cloud motion. The spatial and temporal cloud parameters provided by the method were shown to be useful for interpreting the meteorological processes affecting the development of cloud.

In the framework of the thesis, the data assimilation has been performed using a Kalman filter-like approach. In order to deal with the strong non-linearity of the observation process, an Ensemble Kalman Filter (EnKF, [7]) has been considered. Unfortunately, the lack of time has not allowed its implementation.

### 1.3 Thesis overview

In a first part (Chapter 2), the data used in the framework of the thesis is described. In order to use the data provided in an efficient way, the latter has to be preprocessed (Chapter 3): raw satellite images do not allow the distinction between cloudy and non-cloudy regions. The discrimination relies on a static model introduced in Section 3.1. Subsequently, the clouds will be modelled using a Radial Basis Function network in Chapter 4. Once the model has been set up, its properties have to be investigated: the Kalman filter-like approach undertaken is described and tested in Chapter 5.

## Chapter 2

# Description of the dataset

### 2.1 Acquisition process

The images of the globe that will be used in the project are taken from EUMETSAT's METEOSAT satellites (geostationary orbit, about  $36000km$  above the equator). Currently, EUMETSAT is operating METEOSAT satellites at positions of 0 degrees longitude and at 63 degrees East. The images are taken on a half-hour basis (ten minutes basis for visible images) and after processing at the control centre in Darmstadt, they are re-disseminated via the satellite to the user community (<http://www.eumetsat.de/>).

The imagery itself consists of rectangular arrays of 8-bit image pixels [16]. The satellite is spin-stabilised, and the data are acquired at the rate of one image line per satellite rotation. Each time the satellite rotates, the radiation detectors for the various channels (Visible, Infrared, Water Vapour) pan across the Earth 'horizontally' from east to west, acquiring one line of data. Between each rotation, a stepping tilt mirror in the camera optics adjusts position so that the next acquired line is offset 'vertically' from the last.

The baseline for the Rapid Scanning Service (RSS) is quasi-permanent rapid scan imaging with METEOSAT-6. This means that rapid scanning is performed for a certain period and then interrupted for a short time before rapid scanning continues once more. The period between rapid scanning sessions is used to perform full Earth scanning, needed to derive navigation information for the image processing system. The Full Rapid Scan region (FRS) scan area for the RSS covers a latitude range from approximately 10 degrees North to 70 degrees North (northern hemisphere). A subset of the FRS region that covers the European sector is chosen for this project (Figure 2.1).

Three different types of images are provided by EUMETSAT: visible images (VIS) (Figure 2.2), infrared images (IR) and water vapour images (WV). The visible images which measure the radiance of the reflected sunlight on the Earth have been used in the framework of the project.



Figure 2.1: European sector.



Figure 2.2: European sector: visible satellite image.

## 2.2 Dataset

The dimensions of the visible images are the following: 800 pixels by 1800 pixels (800 lines by 1800 columns). In order to reduce the amount of information to deal with, the resolution of the images has been reduced by four: their new dimensions are 400x900 pixels.

The working dataset is composed of 570 visible images. Since the modelling of the space and time behaviour of the clouds needs to be performed for a large sector, the images chosen for the dataset have been restricted to a time window for which the European region is illuminated. Therefore the dataset comprises images taken from 09.30 to 13.00 (GMT) for the period extending from the 15<sup>th</sup> of February 2004 to the 14<sup>th</sup> of March 2004.

The set of images used to find the standard reflectance of the ground (Section 3.4) is composed of the images from the 15<sup>th</sup> of February 2004 to the 12<sup>th</sup> of March 2004. The last two days of images will be used to test the cloud forecast effectiveness (see Section 5.3).

## 2.3 Summary

In the present Chapter, the acquisition process of satellite images has been introduced as well as the working dataset comprising of 570 visible images of the European sector, taken from 09.30 to 13.00 (GMT) for the period extending from the 15<sup>th</sup> of February 2004 to the 14<sup>th</sup> of march 2004.

## Chapter 3

# Image preprocessing

Visible satellite images consist of both cloud and ground signals. The conception of a space and time model for clouds relies on relevant features of the atmosphere: shape and brightness of clouds. Thus, to keep proper information about clouds, the ground signal has to be removed from visible satellite images. To do so, a static model (introduced in Section 3.1) has been set up. The notions needed to perform the discrimination between cloudy and non-cloudy regions are then presented in the next Sections.

### 3.1 Observation model

A satellite image can be thought of as the weighted sum of the cloud albedo and the land albedo where the weights are related to the 'cloud optical thickness'. Thus, the intensity,  $I$ , of the satellite image in a specific pixel is given by:

$$\mathbf{I} = S(\mathbf{s}, t)(\alpha(\mathbf{s}, t)R_c + (1 - \alpha(\mathbf{s}, t))R_g(\mathbf{s}, t)),$$

where:

- $S(\mathbf{s}, t)$  is the solar irradiance;
- $\mathbf{s}$  is the space vector ( $\mathbf{s}$  has two components: the longitude and the latitude);
- $t$  is the time index;
- $\alpha(\mathbf{s}, t)$  is the 'percentage of cloudiness' or 'cloud optical thickness';
- $R_c$  is the reflectance of the clouds;
- $R_g(\mathbf{s})$  is the reflectance of the ground.

The discrimination of the cloudy regions from the non-cloudy ones relies on the percentage of cloudiness  $\alpha(\mathbf{s}, t)$  of a pixel: below a given percentage of cloudiness, the pixel will be labelled as non-cloudy. Conversely, above this fixed value, the pixel will be labelled as cloudy.

The ratio of the intensity of the image over the solar irradiance has values in the interval  $[0, 1]$ . The clouds are assumed bright and thus their reflectance  $R_c$  is fixed to 1. Working out the percentage of cloudiness of a given pixel comes down to solving a linear equation having first computed the ratio of the intensity of this pixel over the solar irradiance (Section 3.2) and the standard reflectance of the ground at this location (Section 3.4).

## 3.2 Computation of the solar flux density

The direct solar flux density ( $W.m^{-2}$ ) at a given pixel is computed using the methods introduced in [6] by Dozier *et al.* The computation relies on:

- latitude;
- julian day;
- hour angle (which introduces the longitude).

The julian day is nothing but the number corresponding to the considered day of the year (e.g. the julian day corresponding to the first of January is 1, the one of the last day of the year is either 365 or 366).

To compute the solar flux density of a pixel, the coordinates of the latter have to be expressed in terms of longitude and latitude. At this point, the coordinates of a pixel are expressed in column and line form. A conversion routine linking the digital to geographic coordinates has to be implemented.

## 3.3 Digital to geographic coordinates

In the basic geometry of the Earth as seen from an equatorial geostationary satellite such as METEOSAT, the equator and the sub-spacecraft meridian define planes that pass through the satellite position. In consequence, these lines of latitude and longitude appear as straight lines in the perspective view. As a result of the satellite perspective, all other lines of latitude and longitude appear to bow outwards from these fixed central planes [16].

### 3.3.1 Digital coordinate geometry

Starting with the line coordinate, each line of data is gathered with the tilt mirror fixed at a constant angle with respect to the spin axis of the satellite. This means that each line corresponds to a cone swept out in space. For the middle line, the mirror sweeps around the equatorial plane. For other lines, the mirror is tilted out of this plane. Each successive line is gathered by tilting the mirror by one additional step, with each step corresponding to a fixed angular increment.

For the pixel direction, each pixel in a line represents a fixed fraction of the total cone swept out by that line. The equivalent pixels on successive lines are exactly aligned vertically as seen from the spacecraft.

### 3.3.2 Digital to geographic conversion routine

The basis of the algorithm is a determination of the intersection point between the surface of the Earth and the viewing line from the satellite. There are three possibilities for this intersection:

1. no intersection;
2. one intersection (Figure A.1);
3. two intersections: front/visible and back/invisible sides of the disc.

The viewing lines can be thought of in terms of lines drawn from an origin at the spacecraft through an ‘image plane’ in front of the spacecraft. To determine the intersection point(s), equations must be defined for the viewing line and the Earth’s surface.

The script of the conversion routine is discussed in Appendix A.

### 3.3.3 Test of the conversion routine

In order to check the accuracy of the digital to geographic coordinates conversion routine, the following test has been carried out:

1. the coastline file<sup>1</sup> (Figure 3.1), provided by EUMETSAT, is resized to be of the same size as the images used in the project;
2. the line/pixel number to geodetic coordinates conversion is performed on the coastline file. Hence, at this stage the geodetic coordinates of each pixel of the coastline image are known;
3. using MATLAB’s toolbox M\_MAP, the geodetic coordinates of the coastline are plotted on a projection of the Earth viewed by a geostationary satellite;
4. the layout computed using the conversion routine is consistent with the projection (Figure 3.2).

## 3.4 Computation of the reference brightness

As discussed in Section 3.1, the standard reflectance of the ground at a given pixel has to be computed in order to determine its ‘percentage of cloudiness’  $\alpha(\mathbf{s}, t)$  therefore its nature: cloudy or non-cloudy. The first method considered to compute the reference brightness of the ground was the one introduced in [8] i.e. the computation of the standard brightness as the mean of the first component of a three

---

<sup>1</sup>Layout of the coast lines on a plain background (GIF file)



Figure 3.1: Layout of the coastline of the European region.

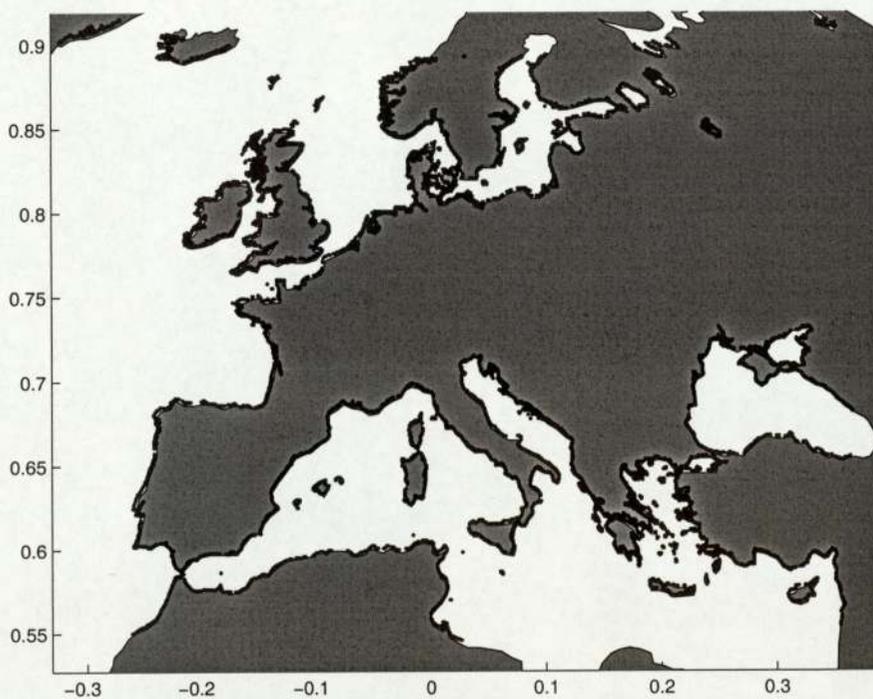


Figure 3.2: Plotting of the geodetic coordinates (black points) of the coastline on a projection of the Earth viewed by a geostationary satellite (detail).

component Gaussian Mixture Model (Section 3.4.1). Nevertheless, the results obtained using this method were not satisfactory and thus a second method has been used: the computation of standard reflectance of the ground as the minimum brightness (Section 3.4.2).

### 3.4.1 Gaussian Mixture Model

The use of a Gaussian Mixture Model to compute the standard reflectance of the ground is motivated by the properties of the histograms of brightness (discussed next paragraph) and the need for a good estimate of the reference albedo. Since the reference brightness of the ground is computed as the mean of the distribution of the lowest brightnesses (that is to say the first component of the Gaussian Mixture Model), the estimate obtained is consistent with the whole dataset.

For some randomly chosen pixels, histograms of brightness over the dataset have been plotted: the latter count the number of occurrences of a given brightness (in the range  $[0, 255]$ ) for the considered pixel. The histograms present common characteristic features:

- peak in brightness for values around 50 and values between 80 and 170;
- Gaussian-like shape of the peaks (Figure 3.3).

Considering that the reflectance of the ground is darker than the reflectance of the clouds and taking into account the features of the brightness histograms, the standard brightness of the ground will be computed as the mean of the first component of a three component Gaussian Mixture Model. The three components are assumed to characterise the brightness of the ground, the brightness of thick clouds (mixing of clouds and ground) and finally the brightness of clouds.

#### Gaussian Mixture Model

In a mixture model, a probability density function is expressed as a linear combination of basis functions. A model with  $M$  components is written in the form:

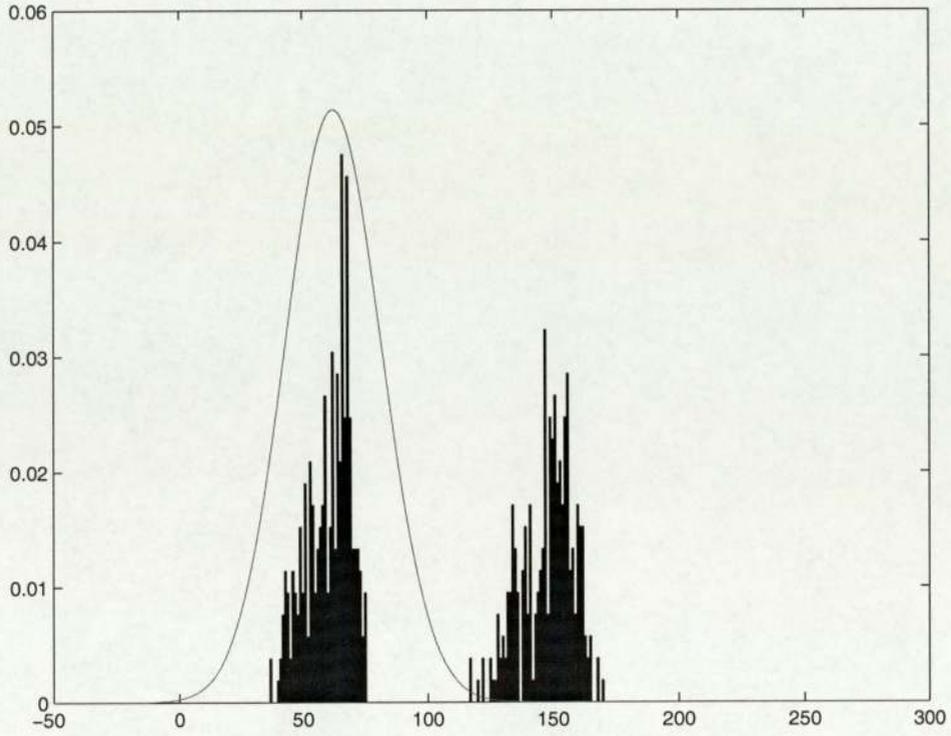
$$p(\mathbf{x}) = \sum_{j=1}^M p(\mathbf{x}|j)P(j).$$

The  $M$  functions  $p(\mathbf{x}|j)$  are the components density functions, with  $\int_{-\infty}^{\infty} p(\mathbf{x}|j)dx = 1$ . The probabilities  $P(j)$  which combine the individual models are the mixing coefficients and,

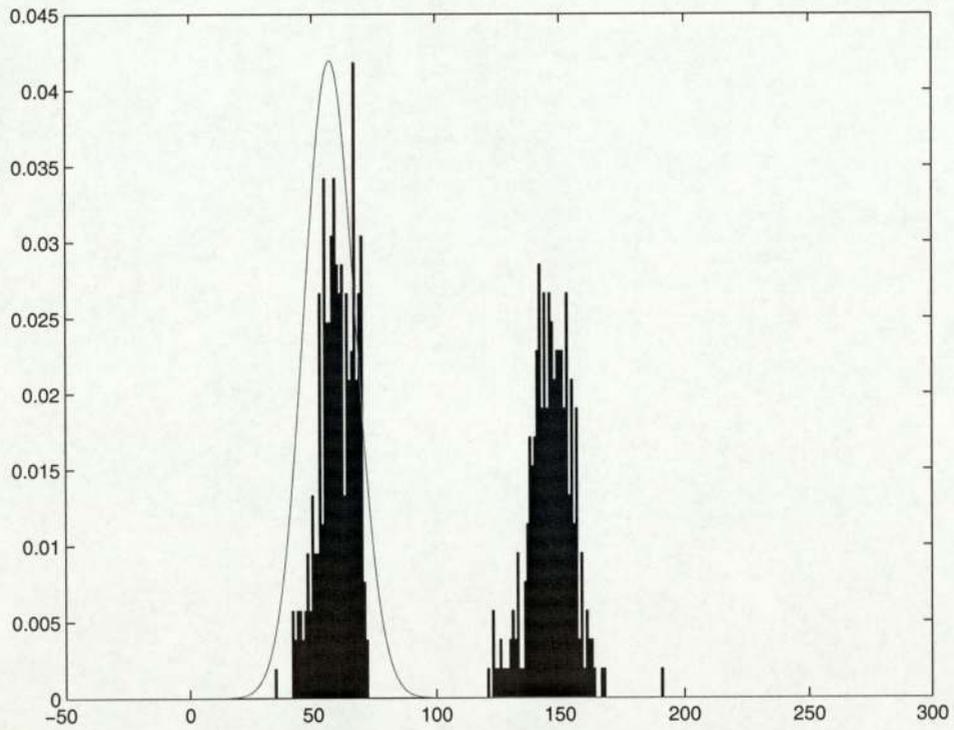
$$\begin{aligned} 0 &\leq P(j) \leq 1, \\ \sum_{j=1}^M P(j) &= 1. \end{aligned}$$

In the case where the component densities are Gaussian (of spherical covariance  $\Sigma_j = \sigma_j^2 \mathbf{I}$ ):

$$p(\mathbf{x}|j) = \frac{1}{(2\pi\sigma_j^2)^{d/2}} \exp \left\{ -\frac{(\mathbf{x} - \mu_j)^2}{2\sigma_j^2} \right\},$$



(a) Pixel of pixel-line number coordinates (181,101).



(b) Pixel-line number coordinates (361,41).

Figure 3.3: Histograms of brightness and first component of a three component Gaussian Mixture Model for two pixels.

In order to determine the parameters in a mixture model, the principle of maximum likelihood is applied. To find the maximum likelihood parameters of the components of the Gaussian Mixture Model a re-estimation algorithm is applied: the EM algorithm (explained in Appendix B).

### Results

The Figure 3.4 is the image obtained using the mean of the first component of a three component Gaussian Mixture Model.



Figure 3.4: Standard reflectance of the ground of the European sector (three component Gaussian mixture model case).

The resulting image has black and white straight lines. The latter are flaws: actually, each pixel appearing as a black or white pixel (and being part of a straight line) is a pixel for which the computation of the mean of the first component of a three component Gaussian Mixture Model has failed.

Some disturbance in the acquisition process of the satellite images is at the origin of the flaws. Unscheduled interruptions of the acquisition process such as mechanical failures lead to the transmission of inaccurate data. Basically, the flaws consist of missing lines of information (the images being acquired on a line by line basis). The presence of such artifacts in the dataset creates a bias: the first component of the three component Gaussian Mixture Model used does not represent the distribution of the brightness of the ground anymore.

The computation of the components of the three components Gaussian Mixture Model being time consuming (around 10000 values seemed to be inaccurate), it has been decided to compute the standard reflectance of the ground as the minimum of the values taken by a given pixel on the whole dataset.

### 3.4.2 Minimum brightness

Computing the reference brightness of the ground as the minimum of the values taken by a given pixel over the whole dataset is a convenient way to discard the inaccurate values (due to disturbances

in the acquisition process). A minimum value, above which the values taken by the brightness of the pixel are assumed to be accurate, is set up. Using such a method leads to the creation of a map of minimum brightnesses (Figure 3.5), unlike the previous method which would have led to the creation of a map of mean minimum brightnesses.



Figure 3.5: Minimum brightness of the ground of the European sector.

### 3.5 Discrimination cloudy/non-cloudy regions

The observation model introduced in Section 3.1 can now be solved:

- $\mathbf{I}$  provides the intensity of a given pixel ;
- $S(\mathbf{s}, t)$  is computed according to the method introduced in Section 3.2 ;
- $R_c$  is set to 1 ;
- $R_g(\mathbf{s})$  is computed using the minimum brightness method (Section 3.4.2).

The percentage of cloudiness  $\alpha(\mathbf{s}, t)$  must be in the range  $[0, 1]$ , so the variable  $\gamma(\mathbf{s}, t)$  has been introduced:

$$\alpha(\mathbf{s}, t) = \frac{1}{1 + \exp(-\gamma(\mathbf{s}, t))}.$$

A threshold is applied on the values of  $\gamma(\mathbf{s}, t)$ : below a given value (set up by a human expert), the pixel considered will be labelled as non-cloudy while above this same value, the pixel will be labelled as cloudy. The value of the threshold depends on the period considered. Actually, the 'length' of the dataset is one month (during which modifications in the albedo, due to vegetation growth or melting of the snow, occur) and the standard reflectance of the ground has been computed as the minimum brightness. Therefore, the estimate of the reflectance of the ground is not as robust as if computed as the mean of the distribution of the brightness of the ground (the first component of the Gaussian Mixture Model presented in Subsection 3.4.1).



(a) Visible satellite image.

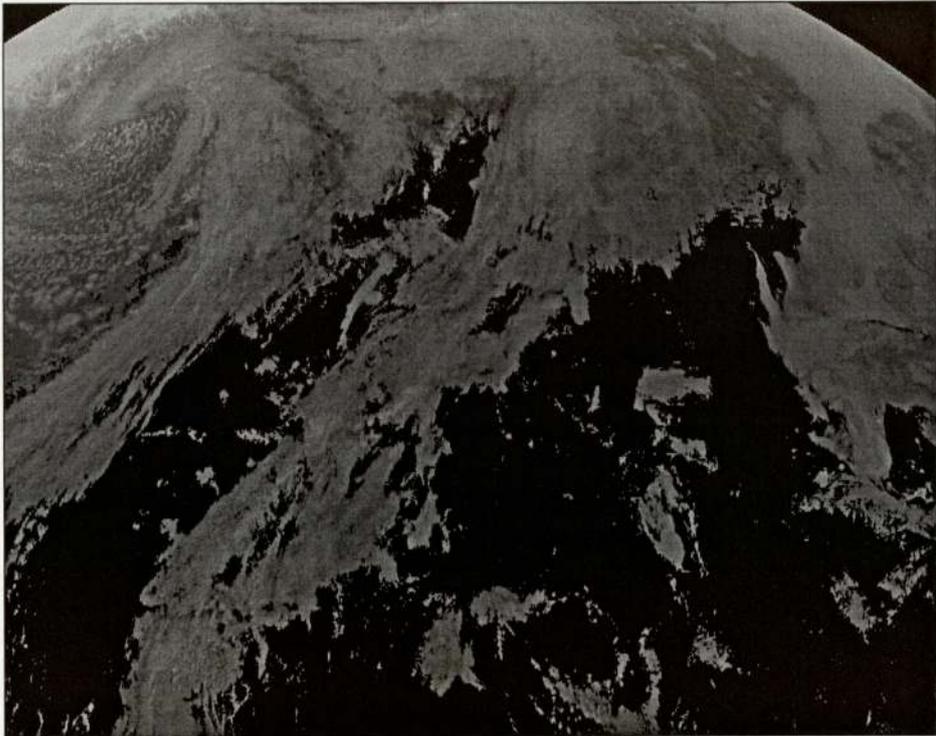


(b) Discriminated image: only the clouds appear ( $\gamma$  representation).

Figure 3.6: Example of discrimination cloudy/non-cloudy regions for the 13<sup>th</sup> of March 2004 at 12.00 (GMT).



(a) Visible satellite image.



(b) Discriminated image: only the clouds appear ( $\gamma$  representation).

Figure 3.7: Example of discrimination cloudy/non-cloudy regions for the 14<sup>th</sup> of March 2004 at 10.00 (GMT).

### 3.6 Discussion

The Figures 3.6 and 3.7 are composed of two sub-figures:

**Figure 3.6(a) and Figure 3.7(a)** are the original raw visible satellite images;

**Figure 3.6(b) and Figure 3.7(b)** are the corresponding resulting images obtained after the discrimination between cloudy and non-cloudy regions has been performed ( $\gamma$  representation).

The format of the original raw visible satellite images and the one of the corresponding resulting images differ: the former have been generated using the IMAGE PROCESSING TOOLBOX function `imshow`, the latter have been generated using the MATLAB function `image`. However, both images have the same size: 400x900 pixels.

At first sight, the resulting images obtained after the discrimination between cloudy and non-cloudy regions are satisfactory: global shape and brightness of clouds are faithful to the raw visible satellite images. Nevertheless, the information is ‘damaged’ at the borders of the preprocessed images: at the north-west and north-east corners of the resulting images, the shape and brightness of the discriminated clouds are not reliable (particularly in Figure 3.6(b)). The inaccurate representation has various origins:

- the ‘map’ of the standard reflectance of the ground is not robust (due to the ‘minimum brightness’ method applied, introduced in Section 3.4.2). Therefore, the threshold applied on the whole European sector does not lead to an accurate discrimination;
- the albedo at the north-west and north-east corners of the raw visible satellite images is not relevant due to the penetration angle of the sunlight in the atmosphere. No accurate information can be extracted from these parts of the original images;
- different colormaps have been used leading to a loss of information.

Regions where the discrimination between cloudy and non-cloudy regions has been performed accurately have experimentally been determined. In the following Chapters, a sub-region of the European sector will be used to perform an accurate discrimination. The sub-region considered stretches from lines 101 to 400 (from the top of the image) and from columns 151 to 750 which gives a sector of dimensions 300x600 pixels.

#### Figure 3.6

The resulting image after discrimination (Figure 3.6(b)) presents poor results on regions such as west of France, south of England and Central Europe: actually the ground signal has not been completely removed. The threshold applied to the ‘cloud optical thickness’ is responsible for those poor results. Nevertheless, finding the optimal value of this threshold is a tough task: a threshold leading to a satisfactory discrimination on the northern region of the European sector will probably perform

bad on its southern region (as said before, this fact is due to the weak robustness of the ‘minimum brightness’ method applied to compute the standard reflectance of the ground). Moreover, applying a strong threshold completely removing the ground signal will lead to the loss of relevant information about the structure of the clouds. A compromise, estimated by a human expert, is set up between accuracy of discrimination and level of discrimination.

### Figure 3.7

The same comments as for Figure 3.6 apply in this case. A flaw in the acquisition process can be seen on Figures 3.7(a) and 3.7(b): a straight black line appears on both images. After applying the static model described in Section 3.1, the flaw has been kept.

## 3.7 Summary

In the present chapter, the following notions have been introduced:

- the need for an observation model, used to perform the discrimination between cloudy and non-cloudy regions, has been emphasized. The model itself is described in Section 3.1. Likewise, the notions used in the observation model such as the direct solar flux density and the reference brightness of the ground have been described in the following sections;
- the direct solar flux density as introduced by Dozier *et al.* [6] has been described in Section 3.2. The routine used to perform the digital to geographical coordinates conversion, needed to compute the direct solar flux density, has been explained and tested in Section 3.3;
- the computation of the reference brightness of the ground following the Gaussian Mixture Model method (Section 3.4.1) and the minimum brightness one (Section 3.4.2) has been developed in Section 3.4. The reference brightness of the ground has been computed with the minimum brightness method, the Gaussian Mixture Model leading to unsatisfactory results;
- the discrimination between cloudy and non-cloudy is finally performed (Section 3.5). The results lead to think that had the computation of the reference brightness of the ground been performed using the Gaussian Mixture Model method in a proper way, the results of the discrimination would have been more robust.

# Chapter 4

## Modelling of the clouds

The choice of a model is a crucial step in the conception of a space-time model for clouds: the latter has to be complex enough to represent the reality of the physical nature of the clouds but it also has to be flexible enough to allow its manipulation in a convenient way. The existing model of Cornford and Batail ([5] and [3], respectively) is explained in the following paragraphs.

### 4.1 Radial Basis Function network: framework

The clouds can be thought of as groups of cells (also called ‘cloud cells’). Therefore, modelling a cloud comes down to modelling each of its ‘cells’. A convenient way to do so is to use Radial Basis Function networks: each ‘cell’ is modelled as a radial basis function whose parameters (centre, height<sup>1</sup>, width) are determined using an optimisation algorithm.

Radial basis function methods have their origins in techniques for performing exact interpolation of a set of data points in a multi-dimensional space [4]. Considering a mapping from a  $d$ -dimensional input space  $\mathbf{x} = (x_1, \dots, x_d)$  to a  $c$ -dimensional target space  $\mathbf{y} = (y_1, \dots, y_c)$ , a radial basis function network is defined as a weighted sum of  $M$  radial basis functions  $\phi$ :

$$y_k(\mathbf{x}) = \sum_{j=1}^M w_{kj} \phi_j(\mathbf{x}) + w_{k0},$$

where the constant additive terms  $w_{k0}$  are called biases.

Several forms of basis function have been considered, such as the tanh basis function used by Lê [8]. Nevertheless, the Gaussian basis function:

$$\phi_j(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mu_j\|^2}{2\sigma_j^2}\right),$$

has been chosen for the following reasons: the Gaussian is a localized basis function with the property that  $\lim_{|\mathbf{x}| \rightarrow \infty} \phi(\mathbf{x}) = 0$  which can be an interesting property for modelling compact ‘cloud cells’ with clear edges, the number of parameters of the Gaussian basis functions is small. . .

---

<sup>1</sup>the term height does not refer to cloud height but to peak intensity

The parameters of the  $j^{th}$  basis function ( $w_{kj}$ ,  $\mu_j$  and  $\sigma_j^2$ ) have to be optimised in order to have the network as close as possible to the function interpolated. The ‘cloud cells’ are assumed to have a Gaussian-shaped brightness rate (brightest pixel at the centre and normal decay) which motivates the choice of Gaussian radial basis functions.

The cloud field  $C$  being modelled using a RBF network with Gaussian basis functions:

$$C = \sum_{j=1}^M \sum_{k=1}^N h_j \exp \left( -\frac{\|x_k - x_j^c\|^2 + \|y_k - y_j^c\|^2}{2w_j} \right).$$

$C$  can be seen as a mixture of  $M$  Gaussians having centres  $(x_j^c, y_j^c)$ , width  $w_j$  and height  $h_j$ .

The RBF network used in the framework of the thesis is illustrated by the Figure 4.1. The input vectors are made of the geographic coordinates ( the longitude and the latitude) of a considered pixel, whereas the output of the network is its brightness. The number of hidden units depends of the complexity of the image to be modelled as will be seen in Section 4.2.

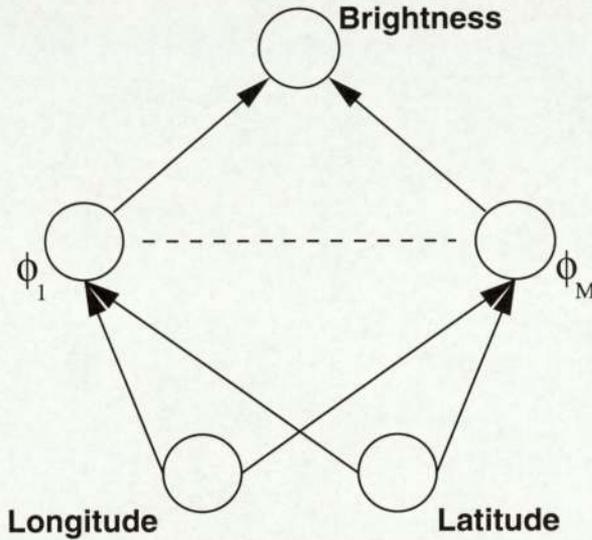


Figure 4.1: RBF network used for the thesis: a cloud is modelled as a mixture of Gaussian radial basis functions.

## 4.2 Network initialisation

The modelling of a cloud as a group of ‘cloud cells’ is a crucial step in the framework of the project. Therefore, the initialisation of the network has to be done with particular care:

1. creation of an empty network  $C_0$  and initialisation of  $I_{temp} = I_0$ , first intensity image in the sequence;
2. search for the brightest pixel in  $I_{temp}$ , of coordinates  $(long_{max}, lat_{max})$ ;
3. modelling of a new ‘cloud cell’: radial basis function of centre  $(x^c, y^c) = (long_{max}, lat_{max})$ , width  $\sigma^2$ , height  $w$ ;

4. local optimisation of the new ‘cloud cell’;
5. removal of the effect of the new ‘cloud cell’ from  $I_{temp}$ ;
6. addition of the new ‘cloud cell’ to  $C_0$ ;
7. loop from step 2 as long as the stopping criterion<sup>2</sup> is not reached.

After having reached the stopping criterion, the global optimisation of the network is performed minimising the negative log-likelihood, using the Scaled Conjugate Gradient algorithm. The local optimisation of the new ‘cloud cells’ has been performed following the same procedure.

### 4.3 Results

The Figure 4.2 is an image obtained after the discrimination cloudy/non-cloudy regions (Section 3.5) has been undertaken on a raw satellite image.

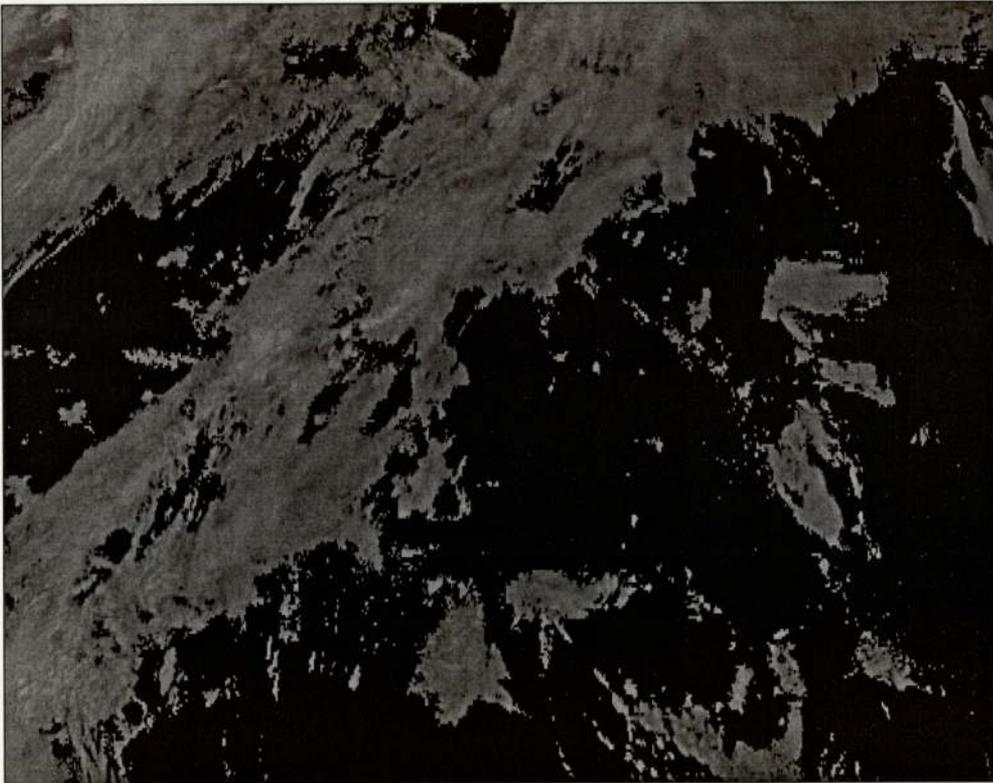


Figure 4.2: European sub-sector:  $\gamma$  values.

Performing the network initialisation on an intensity image of large dimensions (the images used for the project are of size 300 by 600 pixels) can be time-consuming (particularly because of the global optimisation of the network). In order to reduce the computation times, the image is split in smaller images or ‘windows’. The network initialisation algorithm is then run on each ‘window’. The global

<sup>2</sup>minimum brightness or maximum number of cells

optimisation of the network (optimisation of the parameters of the radial basis functions altogether) being performed on a ‘window’ of smaller size, it will be faster than if performed on the whole image. Moreover, one of the original aims of the project was the parallelisation of the model developed by Cornford and Batail ([5],[3]). Splitting the preprocessed image and running the network initialisation algorithm on each ‘window’ is the first step in the parallelisation process, the second one being the parallelisation of the code. However, due to a lack of time, the second step has not been undertaken. For the first step, the ‘windows’ have been delimited with care: defining sharp borders could lead to a poor modelling of the boundary ‘cloud cells’. To avoid the misrepresentation of those cells, each ‘window’ overlaps its neighbours (Figures 4.3 and 4.4): the common regions (whose dimensions are about one fifth of the size of the ‘window’) are weighted using a weight function (Figure 4.5) to prevent the addition of information when reuniting the modelled ‘cloud cells’.

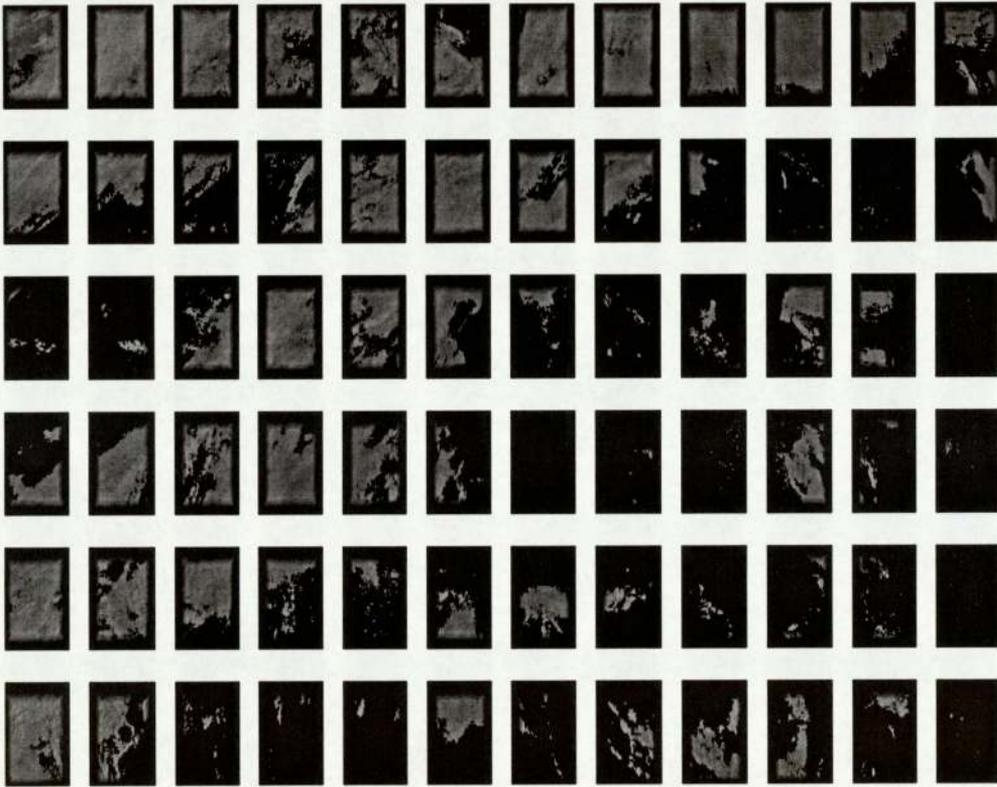


Figure 4.3: Splitting of the original image in windows of size 50x50 pixels.

The split image obtained from the preprocessed image (Figure 4.2) is shown in Figure 4.3. The original length of each basis ‘window’ is 50x50 pixels to which overlaps have been added following the scheme presented in Figure 4.4:

**corner ‘windows’:** overlaps of five pixels length are added to the interior borders (with respect to the preprocessed image) of the ‘window’;

**exterior ‘windows’** (excluding corner ‘windows’): overlap of five pixels length is added to the three interior borders of the ‘window’;

**interior 'windows':** overlaps of five pixels length are added to the borders of the 'window'.

The (intentional) 'loss' of information from the interior of a given 'window' to its boundaries is easily observable in Figure 4.3 as a black frame surrounding the basis 'window'.

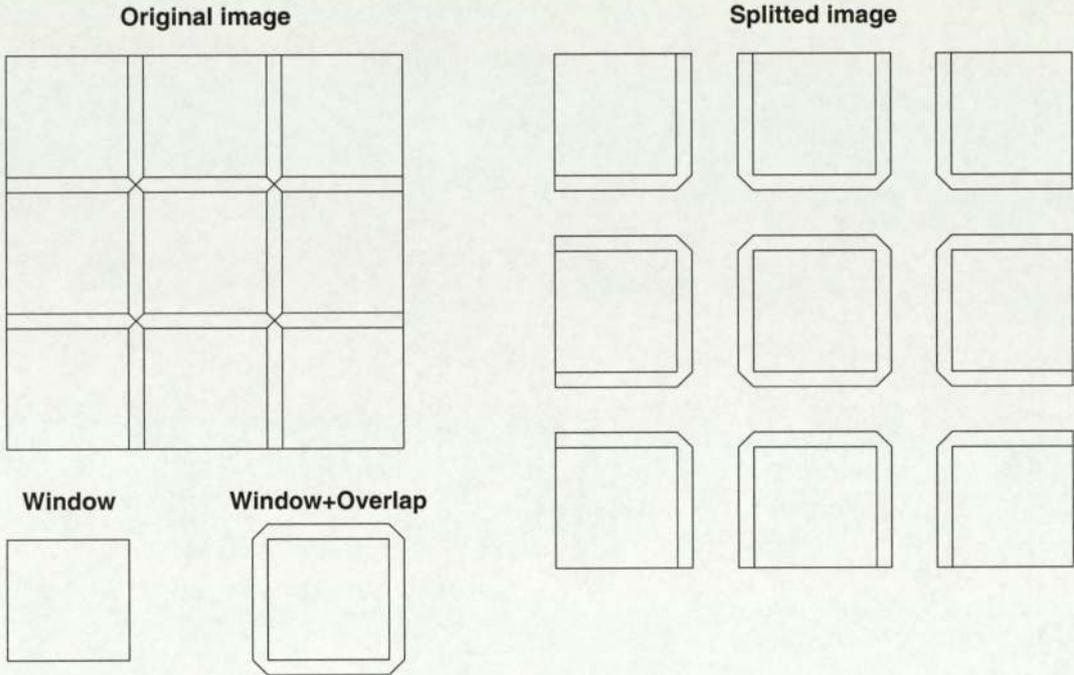


Figure 4.4: Split image and overlap.

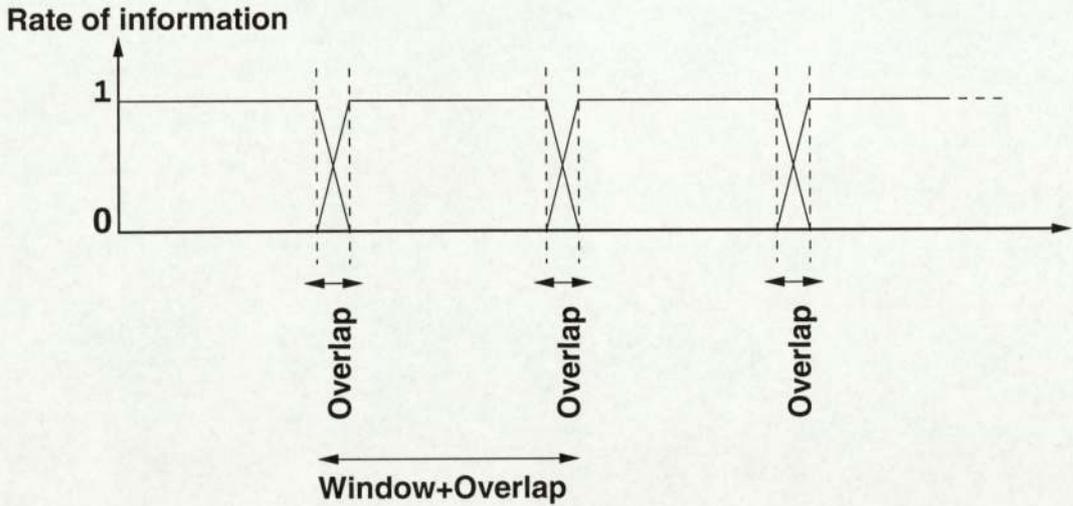


Figure 4.5: Weight function.

The Figure 4.6 is the image obtained after reuniting the modelled ‘cloud cells’ obtained running the initialisation network algorithm on the initial discriminated image (Figure 4.2).

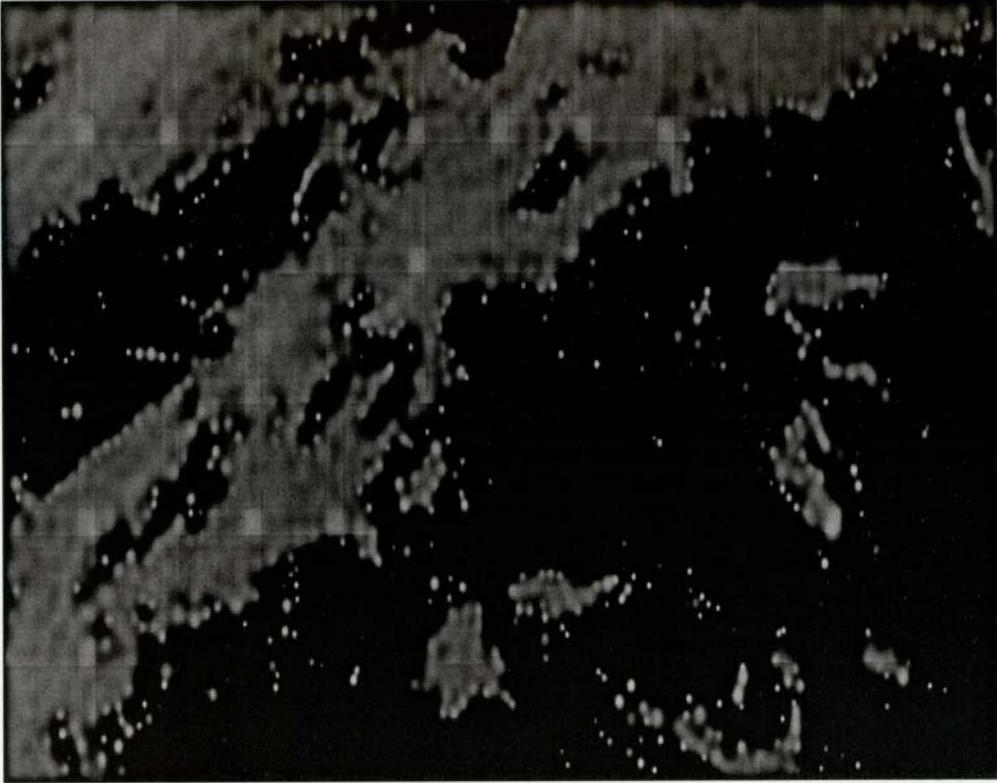


Figure 4.6: Modelled cloud after reunification of the optimised ‘windows’.

The results bring to comments about the display artifacts of the overlapped regions:

1. horizontal and vertical straight lines modelling the boundaries of the overlapped regions;
2. squares of bright intensity (when in a cloudy region) at the intersection of four overlapped regions.

Both of the display artifacts are due to the way of weighting the information of overlapped regions. To explain the origin of the first case two neighbouring ‘windows’ ( $\mathcal{A}$  and  $\mathcal{B}$ ) are considered. In the modelling scheme, the ‘windows’  $\mathcal{A}$  and  $\mathcal{B}$  overlap each other. Hence, they share a common region of overlap,  $\mathcal{S}$ . The outer part of  $\mathcal{S}$  with respect to  $\mathcal{A}$  is the inner part of  $\mathcal{S}$  with respect to  $\mathcal{B}$ . As a weighted function is applied, the outer part of  $\mathcal{S}$  with respect to  $\mathcal{A}$  contains no information whereas the inner part of  $\mathcal{S}$  with respect to  $\mathcal{B}$  contains the whole information. In the outer part of the shared region with respect to  $\mathcal{A}$ , no information will be modelled. In the inner part of  $\mathcal{S}$  with respect to  $\mathcal{B}$ , the ‘cloud cells’ will be modelled using a radial basis function but as the inner part of the shared region is at the border of a region of decay, probably not the whole information of the inner part of the shared region will be modelled. Reuniting the modelled ‘windows’  $\mathcal{A}$  and  $\mathcal{B}$  will not lead to the recovery of the whole information of the discriminated image, which appears as darker straight lines.

The second display problem is caused by the way of applying the weight function at the corners of ‘windows’. Considering four neighbouring ‘windows’ with a common overlapped region  $\mathcal{S}$ , the amount of information of the contribution of each ‘window’ to  $\mathcal{S}$  is not weighted enough before performing the initialisation algorithm on each ‘window’. Hence, the effect of the radial basis functions of the contributions to  $\mathcal{S}$  lead, when reunited, to a rate of information greater than 1.

The purpose of the Figure 4.7 is to show the distribution of the centres of the radial basis functions, after having run the network initialisation, within ‘windows’ of 50x50 pixels (overlap of 5 pixels length).

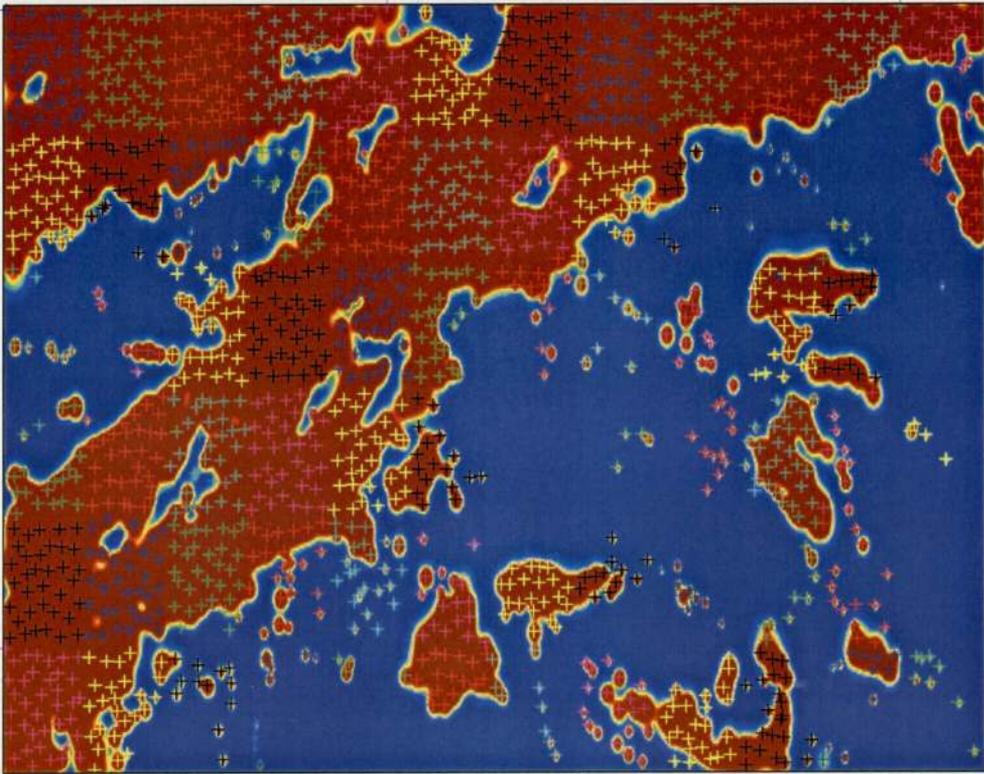


Figure 4.7: Modelled cloud after reunification of the optimised ‘windows’: crosses correspond to the centres of the modelled ‘cloud cells’, different colours correspond to different ‘windows’.

As seen on Figure 4.7, the centres of the modelled ‘cloud cells’ of ‘window’  $\mathcal{A}$  are distinct from the centres of the modelled ‘cloud cells’ of its neighbouring ‘window’,  $\mathcal{B}$  (in the case of an overlapped region shared by two ‘windows’). Hence, the ‘cloud cells’ of the overlapped regions are not modelled twice which is an important result as will be discussed later. Nevertheless, in the case of an overlapped region shared by four ‘windows’, the centres of the modelled ‘cloud cells’ are very close, leading to the following question: “Has an unique ‘cloud cell’ been modelled four times instead of one?”. The examination of Figure 4.6 leads to think that less than four different ‘cloud cells’ should have been modelled for those specific regions. This fact gives away a flaw in the split process or in the reunification process. As the weight function gives good results for overlapped regions shared by two ‘windows’,

the flaw can probably be solved by investigating the way of reducing the amount of information at the corners of neighbouring ‘windows’ sharing a four ‘windows’-overlapped region.

However, the results obtained tend to prove that the split process can be undertaken and leads to satisfactory results if performed properly, as ‘cloud cells’ are modelled once.

An example of the initialisation algorithm performed on a single window can be found in Appendix C.

## 4.4 Summary

In the present Chapter, the following notions have been described:

- considering a cloud as group of elementary ‘cloud cells’ leads to the modelling of the cloud field as a RBF network with Gaussian radial basis functions (Section 4.1);
- the algorithm used to proceed to the initialisation of the network has been introduced in Section 4.2;
- finally, the initialisation algorithm is performed on real data. A new approach of the initialisation is undertaken: the original discriminated image is split into ‘windows’ overlapping each other. After having run the initialisation algorithm on each ‘window’, the ‘windows’ are reunited in a efficient way (despite minor display artifacts, see Section 4.3). The algorithm coupled with the ‘window’ approach seems to perform well.

# Chapter 5

## Dynamics of the model

Modelling the processes at the origin of the motion of the clouds is a complex task. In the framework of the thesis, the *advection field*, which can be thought of as a ‘wind flow’ responsible for the moving of the clouds, has been considered. Other processes, such as phenomena responsible for growth and decay of the clouds, could have been investigated but the scope of the task and the time granted did not allow it.

### 5.1 Advection equation

The advection equation, which links together the cloud rate  $\mathbf{C}$  and the advection field  $\mathbf{v}$ , models the dynamics of the system considered in the thesis. The *advection field* is at the origin of the moving of the ‘cloud cells’ or, as assumed in the framework of the thesis, the centres of the ‘cloud cells’: each whole cell is advected as a block. Thus, the modifications of the characteristics of the ‘cloud cells’, such as their width and height, are not considered during the dynamic step. The *advection equation* is given by:

$$\begin{cases} \frac{\partial \gamma(\mathbf{s}, t)}{\partial t} = -\mathbf{v} \cdot \nabla \gamma + \epsilon, \\ \mathbf{v} = \mathbf{v}(\mathbf{s}, t) = (u, v), \end{cases}$$

where:

- $\gamma(\mathbf{s}, t)$  is the variable related to the cloud field at a given time and location;
- $t$  is the time;
- $(u, v)$  is the advection vector;
- $\epsilon$  is an error term.

#### Advection field $\mathbf{v}$

As said earlier, one the processes responsible for the moving of the clouds, the advection field, can be thought of as a ‘wind flow’. In the model developed in the framework of the thesis, the space of

definition of the advection field is restricted to the ‘cloud cells’ centres, meaning that each cell has its associated movement.

## 5.2 State-space model: Kalman filter

The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modelled system is unknown [18].

The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, the equations for the Kalman filter fall in two groups: *time update* equations and *measurement update* equations. The time update equations are responsible for projecting forward in time the current state and error covariance estimates to obtain the *a priori* estimates for the next time step. The measurement update equations are responsible for the feedback, that is to say for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate.

In the framework of the thesis, the data observed consists in a sequence of visible satellite images. Given a sequence of images up to and including time  $t$ , denoted  $\mathcal{I}_t = (I_1, \dots, I_t)$ , the model should be able to predict the state at time  $t + 1$ ,  $I_{t+1}$ .

### Measurement or observation equation:

For a discrete linear model, the *measurement* equation is as follows:

$$I_t = F_t \theta_t + \epsilon_t^I,$$

where  $\epsilon_t^I$  is a Gaussian noise term with mean zero and covariance matrix  $\Sigma_t^\epsilon$ .

For a non-linear model, the *measurement* equation becomes:

$$I_t = F(\theta_t) + \epsilon_t^I.$$

### State or transition equation:

The state vector can evolve between timesteps according to the *state* equation:

$$\theta_{t+1} = E_t \theta_t + W_t \nu_t$$

$\nu_t$  is a white noise term of covariance matrix  $\Sigma_t^\nu$ .

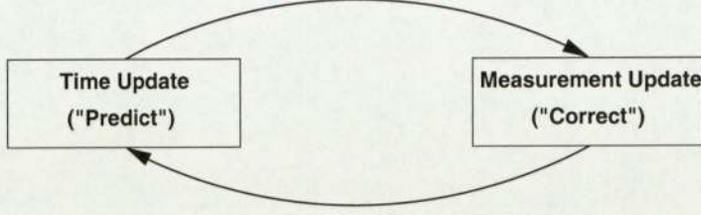


Figure 5.1: The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate by an actual measurement at that time.

### 5.2.1 Evolution of $\gamma$

The evolution equation, as described in the Kalman filter-like approach, used to make evolve the variable  $\gamma$ , linked to the cloud rate  $C$ , is the advection equation:

$$\frac{\partial \gamma(\mathbf{s}, t)}{\partial t} = -\mathbf{v} \cdot \nabla \gamma + \epsilon,$$

which, after discretisation, yields the following difference equation:

$$\gamma_{t+1}(u, v) = \gamma_t(u, v) - \mathbf{v} \Delta t \nabla \gamma.$$

Assuming that the advection field  $\mathbf{v}$  is locally constant because of the slow motion of the clouds, the evolution of the centres of the ‘cloud cells’ is approximated by the following equations:

$$\begin{aligned} x_{t+1|t}^c &\approx x_t^c + u(x_t^c, y_t^c) \Delta t + \epsilon_c, \\ y_{t+1|t}^c &\approx y_t^c + v(x_t^c, y_t^c) \Delta t + \epsilon_c. \end{aligned}$$

Assuming a zero mean Gaussian density of covariance  $\Sigma_{\epsilon_c}$  for the evolution of the noise, the distribution of the estimated centres becomes  $p(x_{t+1}^c, y_{t+1}^c) = \mathcal{N}((\bar{x}_{t+1|t}^c, \bar{y}_{t+1|t}^c), \Sigma_{t+1|t}^c)$ , where:

$$\begin{aligned} \bar{x}_{t+1|t}^c &= \bar{x}_t^c + u_t(\bar{x}_t^c, \bar{y}_t^c) \Delta t, \\ \bar{y}_{t+1|t}^c &= \bar{y}_t^c + v_t(\bar{x}_t^c, \bar{y}_t^c) \Delta t, \\ \Sigma_{t+1|t}^c &= \Sigma_t^c + (\Delta t)^2 \Sigma_{\mathbf{v}_t} + \Sigma_{\epsilon_c}. \end{aligned}$$

### 5.2.2 Evolution of $\mathbf{v}$

The evolution of  $\mathbf{v}$  is assumed to be slow compared to the evolution of the cloud rate  $C$ . In the framework of the thesis, a simple model has been chosen: the white noise  $\epsilon_{\mathbf{v}}$  of covariance matrix  $\Sigma_{\epsilon_{\mathbf{v}}}$  will be added to the advection field to make it change slowly:

$$\mathbf{v}_{t+1|t} = \mathbf{v}_t + \epsilon_{\mathbf{v}}.$$

So, the distribution of  $\mathbf{v}$  becomes  $p(\mathbf{v}_{t+1|t}) = \mathcal{N}(\bar{\mathbf{v}}_t, \Sigma_{\mathbf{v}_t} + \Sigma_{\epsilon_{\mathbf{v}}})$ .

### 5.2.3 Assimilation of $\gamma$

Once the evolution of  $\gamma$  has been performed following the *evolution* equation, and the next observation is available, the estimate has to be corrected by assimilating the information of the new data into the

model. This step requires the approximation of the distribution of  $(\mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h})$  given the new image, which is computed using Bayes' rule.

$$p(\mathbf{x}_{t+1}^c, \mathbf{y}_{t+1}^c, \mathbf{w}_{t+1}, \mathbf{h}_{t+1} | I_{t+1}) = \frac{p(I_{t+1} | \mathbf{x}_{t+1}^c, \mathbf{y}_{t+1}^c, \mathbf{w}_{t+1}, \mathbf{h}_{t+1}) p(\mathbf{x}_{t+1}^c, \mathbf{y}_{t+1}^c, \mathbf{w}_{t+1}, \mathbf{h}_{t+1})}{p(I_{t+1})}$$

Removing the time indexes:

$$p(\mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h} | I) = \frac{p(I | \mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}) p(\mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h})}{p(I)}$$

The approximation of the parameters is performed minimising the error measure function defined by the negative log-posterior:

$$E \propto -\ln p(I | \mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}) - \ln p(\mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h})$$

The posterior distribution,  $p(\mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h} | I)$ , being approximated by a Gaussian distribution (using a Laplace approximation, detailed later), the prior over the parameters is also a Gaussian.

Assuming uncorrelated observations, the negative log-likelihood becomes:

$$-\ln p(I | \mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}) = -\sum_{j=1}^M \ln p(I_j | \mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}).$$

Assuming a Gaussian distribution for the measurement process (satellite) with noise  $\sigma^2$ , the likelihood can be written as a noisy realisation of the model:

$$p(I_j | \mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(I_j - C_j)^2}{2\sigma^2}\right),$$

where  $C_j$  is the model evaluation at point  $(x_j, y_j)$ :

$$C_j = C(x_j, y_j) = \sum_{k=1}^N h_k \exp\left(-\frac{\|x_k - x_j^c\|^2 + \|y_k - y_j^c\|^2}{2w_j}\right).$$

Hence, the negative log-likelihood becomes:

$$-\ln p(I | \mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}) \propto \frac{1}{2\sigma^2} \sum_{j=1}^M (C_j - I_j)^2.$$

The negative log-prior is:

$$\begin{aligned} -\ln p(\mathbf{x}^c, \mathbf{y}^c, \mathbf{w}, \mathbf{h}) &= -\ln p(\mathbf{x}^c, \mathbf{y}^c) - \ln p(\mathbf{w}) - \ln p(\mathbf{h}) \\ &\propto \frac{1}{2} \mathbf{c}^T (\Sigma^c)^{-1} \mathbf{c} + \frac{1}{2} \mathbf{w}^T (\Sigma^w)^{-1} \mathbf{w} + \frac{1}{2} \mathbf{h}^T (\Sigma^h)^{-1} \mathbf{h} \end{aligned}$$

The parameters which minimise the negative log-posterior, called the *maximum a posteriori* parameters, are the parameters for which the posterior distribution is maximised. The posterior distribution obtained being often complex and difficult to integrate, a Laplace approximation is performed.

### Laplace approximation

The Laplace approximation consists in approximating a complex probability distribution  $P$  about its maximum with a more tractable Gaussian distribution  $Q$ . It uses the fact that at the maximum  $\theta_m$ ,  $\ln p(\theta)$  can be approximated using a Taylor expansion:

$$-\ln p(\theta) \approx -\ln p(\theta_m) + \frac{1}{2}(\theta - \theta_m)^T \mathbf{H}(\theta - \theta_m),$$

where

$$\mathbf{H} = \nabla \nabla \ln p(\theta)|_{\theta_m},$$

is the local Hessian matrix. This local expansion defines a Gaussian approximation:

$$q(\theta) \approx \frac{|\mathbf{H}|^{1/2}}{(2\pi)^{k/2}} \exp \left\{ \frac{1}{2}(\theta - \theta_m)^T \mathbf{H}(\theta - \theta_m) \right\}.$$

In the framework of the thesis, the Hessian of the posterior distribution is computed with respect to the whole optimised parameter  $\theta_m$ . Therefore, the posterior distributions of the parameters can be updated using the optimised values of  $\mathbf{x}^c$ ,  $\mathbf{y}^c$ ,  $\mathbf{w}$ ,  $\mathbf{h}$  for the mean vectors and the appropriate parts of the Hessian diagonal for the covariance matrixes.

#### 5.2.4 Assimilation of $\mathbf{v}$

As most of the difficulty of the problem has been put in the evolution of the cloud rate  $C$ , the evolution of the advection field  $\mathbf{v}$  is relatively easy. The linear update of the advection field only relies on the uncertainty over the centres of the cloud rate  $C$ :

$$\begin{aligned} p(\mathbf{v}_{t+1}|I_t, I_{t+1}) &= \int \int p(\mathbf{v}_{t+1}, \mathbf{c}_{t+1}, \mathbf{c}_t | I_t, I_{t+1}) d\mathbf{c}_{t+1} d\mathbf{c}_t \\ &= \int \int p(\mathbf{v}_{t+1}, \mathbf{c}_{t+1}, \mathbf{c}_t) p(\mathbf{c}_t | I_t) p(\mathbf{c}_{t+1} | I_{t+1}) d\mathbf{c}_t d\mathbf{c}_{t+1}. \end{aligned}$$

### 5.3 Running the space and time model of clouds

The purpose of the present Section is to evaluate the performances of the space and time model for clouds. The dataset used to perform experiments on the model consists of two days of visible satellite images as described in Section 2.2: European sector on the 13<sup>th</sup> and 14<sup>th</sup> of March 2004, between 09.30 and 13.00 GMT (images acquired every ten minutes).

#### 5.3.1 Initialisation of the model

Before running the model, the latter has to be initialised.

### Priors

As the model is fully defined by the parameters of the cloud rate  $C$ :  $(\mathbf{x}^c, \mathbf{y}^c, \mathbf{h}, \mathbf{w})$  and those of the advection field  $\mathbf{v}$ :  $(\bar{\mathbf{v}}, \Sigma_{\mathbf{v}})$ , the prior distributions of the latter have to be specified with care: they actually reflect the expectations for their probability densities.

- the cell centres  $\mathbf{x}^c$  and  $\mathbf{y}^c$  are chosen to be Gaussian distributed with mean vector  $\bar{\mathbf{x}}^c = (x^{\bar{c}_1}, \dots, x^{\bar{c}_1})$  and  $\bar{\mathbf{y}}^c = (y^{\bar{c}_1}, \dots, y^{\bar{c}_1})$  (where  $N$  is the number of ‘cloud cells’) and diagonal covariance matrix  $\Sigma^c$ ;
- the heights  $\mathbf{h}$  and the widths  $\mathbf{w}$  are chosen to be log-Gaussian with mean vector  $\bar{\mathbf{h}}$  and  $\bar{\mathbf{w}}$  respectively and diagonal covariance matrix  $\Sigma^h$  and  $\Sigma^w$  respectively;
- $\mathbf{v}$  is defined as a Gaussian Process with mean  $\bar{\mathbf{v}}$  and variance  $\Sigma_{\mathbf{v}}$ .

### Initialisation of $C$

The initialisation of the cloud field  $C$  is performed as described in Section 4.2.

### Initialisation of $\mathbf{v}$

The initialisation of the advection field relies on the two first cloud rates  $C_0$  and  $C_1$ . The latter have been initialised following the initialisation of the network presented in Section 4.2. Then, using the advection equation  $\frac{\partial \gamma(\mathbf{s}, t)}{\partial t} = -\mathbf{v} \cdot \nabla \gamma$  with  $\frac{\partial \gamma(\mathbf{s}, t)}{\partial t} = C_1 - C_0$  and with  $\nabla \gamma$  computed from  $C_0$  using the partial differences between pixels as  $\frac{\partial \gamma(\mathbf{s}, t)}{\partial x}$  and  $\frac{\partial \gamma(\mathbf{s}, t)}{\partial y}$ , the initial value of  $\mathbf{v}$  is computed.

### 5.3.2 Forecast

As soon as the model has been initialised, it can run iteratively, using a sequence of visible satellite images. The Kalman filter-like approach on which relies the model consists of evolution and update steps successively:

1. forecast of the cloud rate  $C$ ;
2. forecast of the advection field  $\mathbf{v}$ ;
3. update of  $C$ ;
4. update of  $\mathbf{v}$ .

Following this scheme, the model can run as long as input data is provided.  $n$  evolution or forecast steps from time  $t$  will lead to forecast time  $t + n$ .

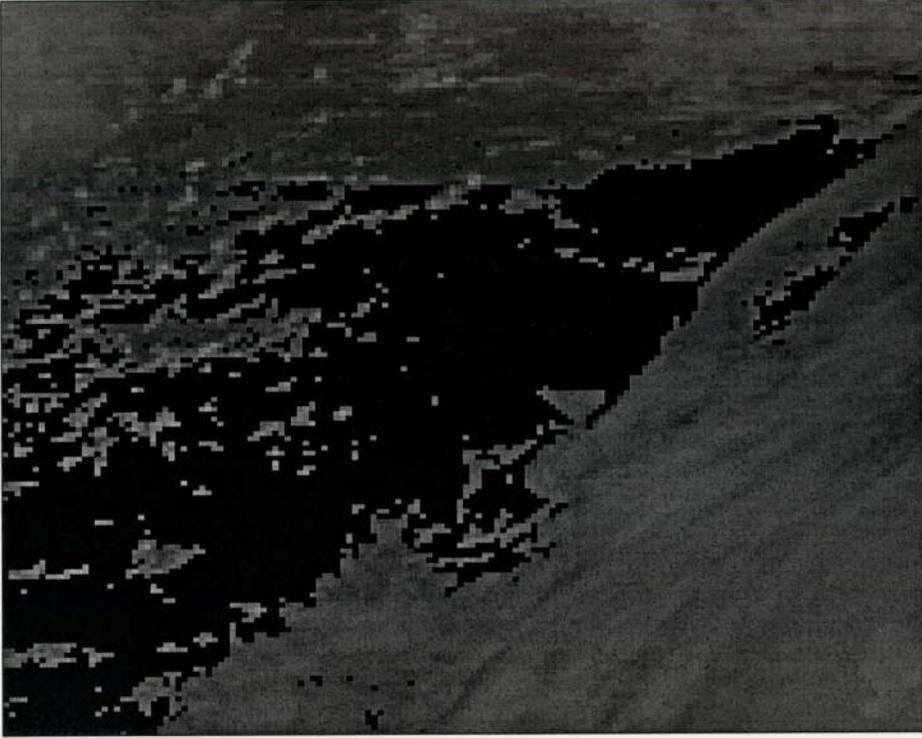
### 5.3.3 Results

In the present case, the dataset used to evaluate the performances of the model comprises of two sets of twenty-two visible satellite images (corresponding to the 13<sup>th</sup> and 14<sup>th</sup> of March 2004). The aim of this section is to prove the accuracy of the model described in Chapter 4: each image is split into ‘windows’; then, each forecast image (such as the Figure 5.4(a)) is a reunited figure where the forecast has been performed on each ‘window’ rather than on the whole image. This approach was motivated by the initial wish of speeding-up the model by distributing the computation using the cluster of the NCRG group. Nevertheless, due to the lack of time, the parallelisation has not been performed although the framework has been developed.

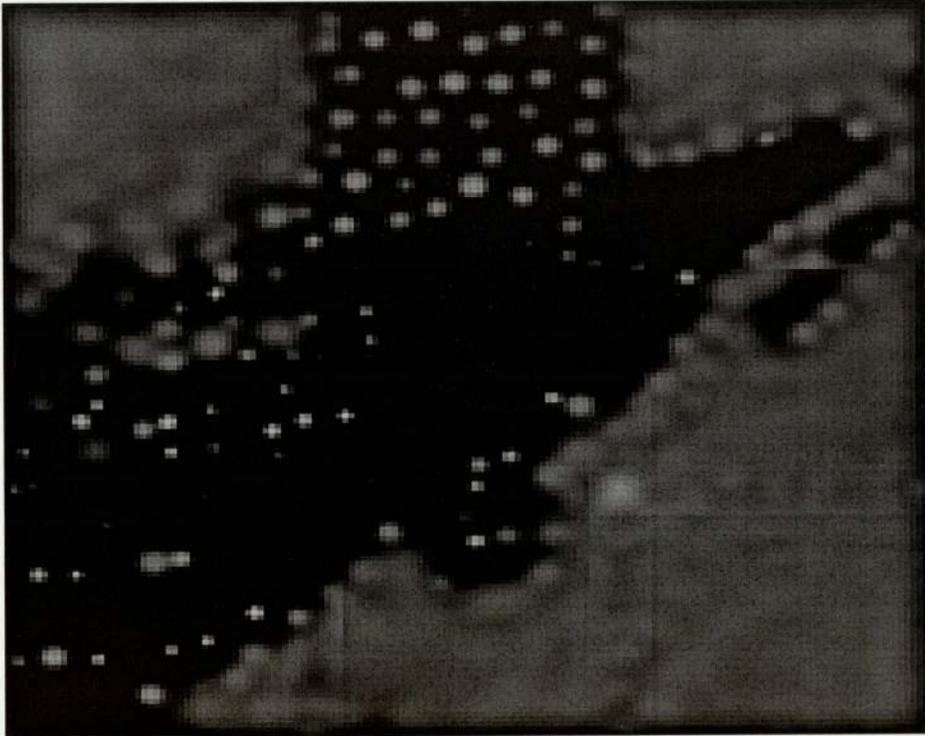
In the following example, a sequence of six images (13<sup>th</sup> of march 2004 from 09.30 to 10.20 (GMT)) is used. In a first time, the model has to be initialised. As described in Section 5.3.1, the initialisation of the advection field  $\mathbf{v}$  requires the two first cloud rates  $C_0$  and  $C_1$ . Hence, the two first images of the sequence will be used to initialise the model (Figures 5.2 and 5.3). Then, the model runs iteratively: the forecast of the model is followed by its update and so on. The Figures 5.4, 5.5 and 5.6, which describe the process, are made up of two sub-figures:

- the first sub-figure (Sub-figures 5.4(a), 5.5(a) and 5.6(a)) corresponds to the forecast of the model;
- the second sub-figure (Sub-figures 5.4(b), 5.5(b) and 5.6(b)) corresponds to the actual state of the model.

The Figures 5.2, 5.3, 5.4, 5.5 and 5.6 are discussed later in the Section.



(a) Actual state ( $\gamma$  representation).



(b) Initialisation of the network: first step.

Figure 5.2: Actual state of the process at time  $t$  and initialisation of the model (first step).

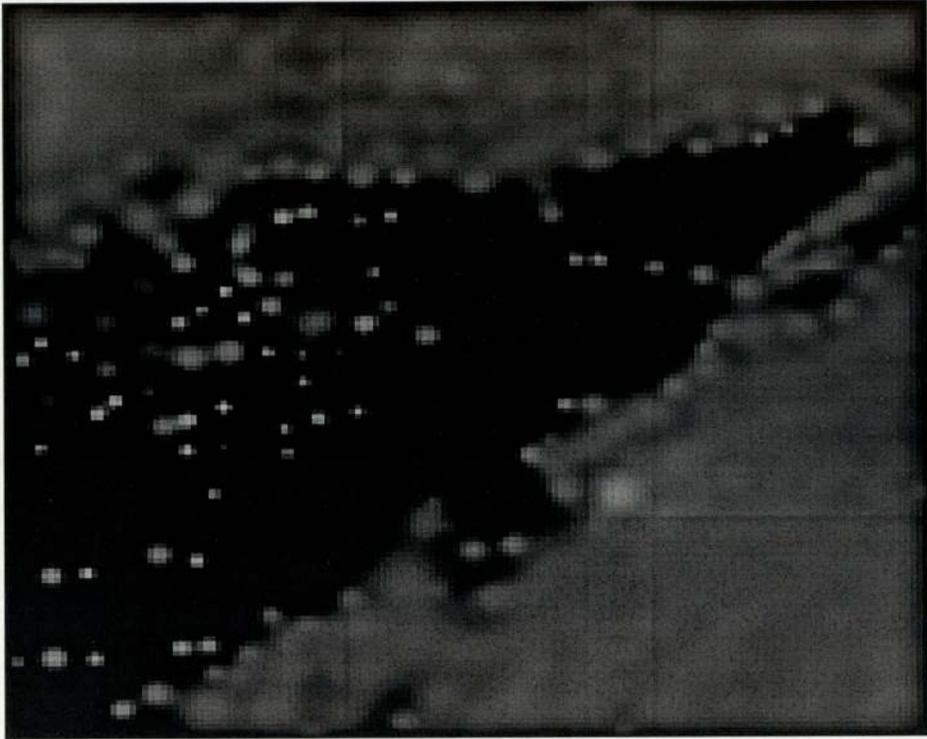


(a) Actual state ( $\gamma$  representation).



(b) Initialisation of the network: second step.

Figure 5.3: Actual state of the process at time  $t + 1$  and initialisation of the model (second step).

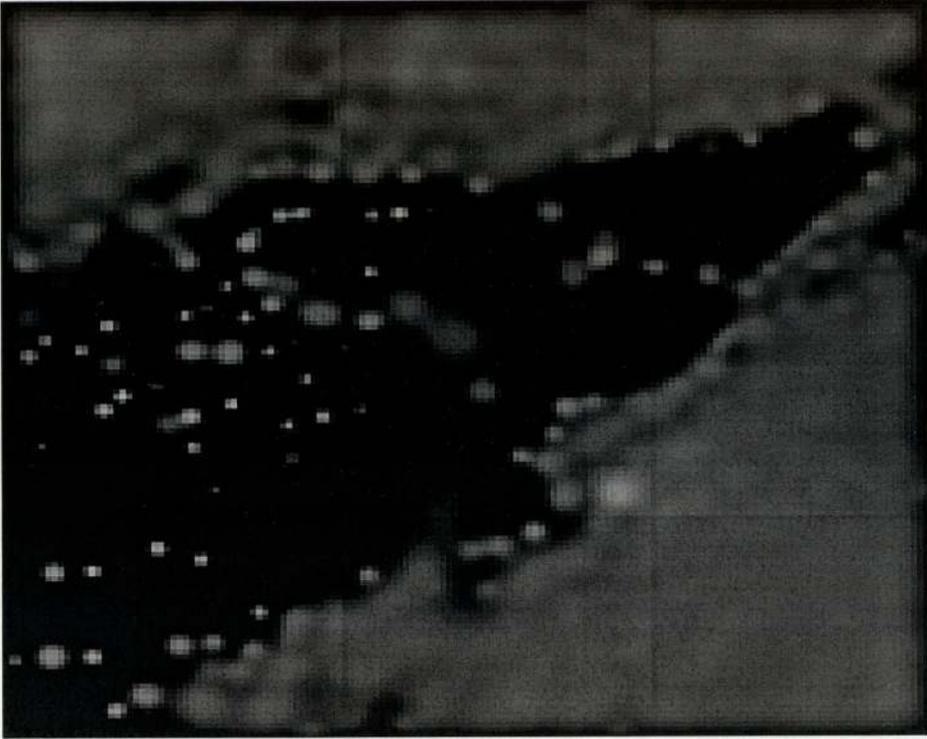


(a) Forecast: second step.



(b) Actual state ( $\gamma$  representation).

Figure 5.4: Forecast of the process at time  $t + 3$  and actual state of the process at time  $t + 3$ .

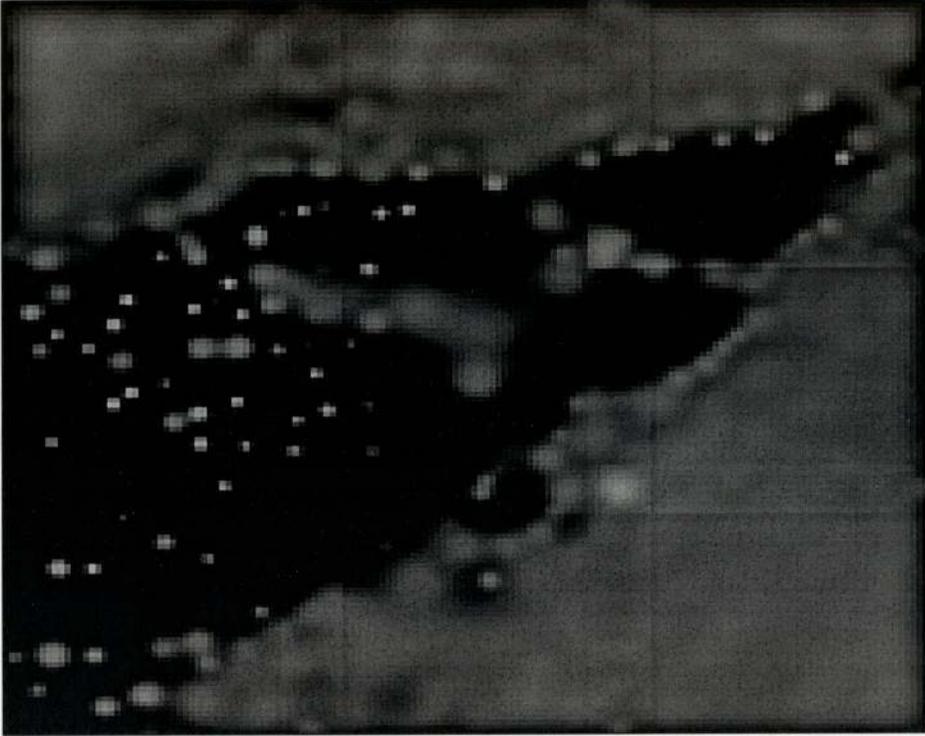


(a) Forecast: third step.



(b) Actual state ( $\gamma$  representation).

Figure 5.5: Forecast of the process at time  $t + 4$  and actual state of the process at time  $t + 4$ .



(a) Forecast: fourth step.



(b) Actual state ( $\gamma$  representation).

Figure 5.6: Forecast of the process at time  $t + 5$  and actual state of the process at time  $t + 5$ .

## 5.4 Discussion

The Sub-figure 5.2(a) is the actual state of the process at time  $t$ , obtained from the discrimination between cloudy and non-cloudy regions performed on a sub-region of the European sector for the 13<sup>th</sup> of march 2004 at 09.30 (GMT). The size of the image is 150 by 150 pixels<sup>1</sup>.

The Sub-figure 5.2(b) is the image obtained after the initialisation of the model at time  $t$  (as described in Section 5.3.1) has been performed. The image is made up of nine elementary ‘windows’ as described in Section 4.3. Initialising the model comes down to initialising each ‘window’, after what the ‘windows’ are reunited and the first ‘cloud rate’ is computed. It is important to notice the display artifact located at the north of the considered sector. This artifact may be due to various factors such as the bad tuning of the prior distributions or the shape of the discriminated clouds of this particular ‘window’.

Except from the usual display artifacts (discussed in Section 4.3) and the display artifact located at the north of the considered sector, the initialisation of the network performs well.

### Comments on Figure 5.3

The Sub-figure 5.3(a) is the actual state of the process at time  $t + 1$ .

The Sub-figure 5.3(b) is the image obtained after the initialisation of the model at time  $t + 1$  has been performed. The same comments as for the initialisation of the network at time  $t$  apply: the display artifact located at the north of the considered sector is still present, leading to the conclusion that it is caused by the bad initialisation of the prior distributions of the parameters of the model. However, the second ‘cloud rate’,  $C_1$ , is computed. The values of  $C_0$  and  $C_1$  can now be used to compute the initial value of the advection field  $\mathbf{v}$  as described in Section 5.3.1.

Again, the results provided by the network initialisation seem quite accurate.

### Comments on Figure 5.4

Now that the model has completely been initialised, the forecasting can be undertaken. The Sub-figure 5.4(a) is the forecast of the model at time  $t + 3$  (the first forecast of the model has not been shown). This estimate of the state of the model at time  $t + 3$  has to be compared to the actual state of the process shown in the Sub-figure 5.4(b).

The first update of the model has allowed the removing of the display artifact discussed earlier. Nevertheless, the usual display artifacts are still present. This fact is caused by the way of computing the forecast image: actually, the forecast step is performed on each ‘window’ rather than on the whole image. This way of proceeding is due to the initial conception of the project which was the parallelisation of the model introduced by Cornford and Batail ([5],[3]). As the parallelisation has

---

<sup>1</sup>the Figures 5.2 to 5.6 have the same format

not been undertaken, the way of computing the forecast image is a limitation to the accuracy of the results. Nevertheless, the latter are satisfactory compared to the actual state of the process.

### Comments on Figures 5.5 and 5.6

The forecast of the model at time  $t + 4$  (Sub-figure 5.5(a)) still performs well by comparison to the actual state of the process (Sub-figure 5.5(b)). The same observation can be made for the forecast of the model at time  $t + 5$  (Figure 5.6). Nevertheless, as a ‘cloud cell’ grows in an overlapped region shared by four neighbouring ‘windows’, the tendency of those regions to be over-represented, as said in Section 4.3, is evident (particularly in Sub-figure 5.6(a)).

## 5.5 Summary

In the present Chapter, the following notions have been developed:

- the advection equation, which models the dynamics of the system, is introduced in Section 5.1 along with the advection field;
- the Kalman filter-like approach used to make the model evolves in time is explained in great details in Section 5.2;
- finally, the results obtained running the model are discussed in Section 5.3. The experiment of the accuracy of the model has been undertaken for a sequence of only six images (which corresponds to a sequence of fifty minutes). As the datasets available for the experiments are made up of twenty-two images (three hours and a half length), the next step would have been to test the model on one of the two datasets. Moreover, the size of the sector studied should have been increased to estimate the accuracy of the model on a high space and time resolutions.

# Chapter 6

## Conclusion

### 6.1 Summary of the work done

The purpose of this section is to account for the work that has been done in 9 months of research as well as the difficulties encountered.

A static observation model has been set up to discriminate raw satellite images into cloudy and non-cloudy regions. The concepts needed to perform the discrimination such as the standard reflectance of the ground, the cloud optical thickness or the digital to geographic coordinates conversion routine have been developed in great detail. Working with satellite images has been a difficult and time-consuming task: computing the standard reflectance of the ground as the mean of the first component of a three component Gaussian Mixture Model has not provided satisfactory results in a reasonable amount of time, leading to the computation of the standard reflectance of the ground as the minimum brightness over the working dataset. The latter method will perform a less robust discrimination than the one performed using the former method. However, the work undertaken for the discrimination is an interesting track for future development as is the ‘window’ approach used in the modelling of the clouds. Once the model has been introduced, it has been applied to real data.

The initial goal of the project was to run the model on the cluster of the NCRG group to speed up the computation times by distributing the tasks. Unfortunately, due to a lack of time, the parallelisation of the model has not been performed and the cluster has only been used in a serial way, lightly speeding up the computation times.

### 6.2 Results

The results have been presented in Sections 3.5, 4.3 and 5.3. Their strengths and weaknesses will be discussed in detail:

**Preprocessing:** The results of the preprocessing step are satisfactory: the discrimination between cloudy and non-cloudy regions is quite good as would expect a human expert. Nevertheless,

work needs to be undertaken in order to compute the reference brightness of the ground using the Gaussian Mixture Model method introduced in Section 3.4.1. The latter will provide a robust computation of the standard reflectance of the ground which will lead to the accurate discrimination of the cloudy and non-cloudy regions using a static threshold;

**Modelling:** The modelling of the clouds using a RBF network provides good results from the point of view of a human expert. The split method introduced in Section 4.3 could prove its robustness if used in a parallel framework. However, the flaws at the origin of the display artifacts should be corrected in order to work on proper modelled cloud fields. Moreover, the parallelisation of the code of the model should be undertaken to test it at its full capacity;

**Dynamics:** The sequence of satellite images used to test the dynamics of the model provided good results. Nevertheless, the latter are inconclusive: the space and time scales are not sufficient to evaluate the long-term performances of the model. A test, based on one of the two datasets of twenty-two images, could be undertaken to evaluate the behaviour of the model on a time scale of three and a half hours.

Improving the results provided by the preprocessing step will have an effect on the modelling step. Fixing the display artifacts occurring during the modelling will provide a more accurate modelling of the cloud fields. Then, the last step will be the tuning of the parameters of the model in a efficient way. Hence, more work can be undertaken to improve the model as will be seen in the next Section.

### 6.3 Further work

In order to perform a robust discrimination between cloudy and non-cloudy regions, the standard reflectance of the ground should be computed using the three component Gaussian Mixture Model method introduced in Section 3.4. The “failure” of the latter method was due to the computation times: computing the parameters of a three component Gaussian Mixture Model for 360000 pixels (the images used in for the project being composed of 400 by 900 pixels) is time-consuming. At the time when the computation had to be performed, the cluster of the NCRG group was not available yet. The use of the cluster would have simplified this task.

The prior distributions of the parameters of the model have to be chosen with care: the model used having been designed for the short-term rainfall forecasting, the prior distributions do not fit to the short-term forecasting of the motion of the clouds. Some of the prior distributions have been tuned to the cloud application problem, nevertheless, due to the lack of time, no deep investigation, supported by tests, has been undertaken. A poor prior distribution of the parameters of the model could be at the origin of artifacts such as the one seen in Figure 5.2(b), lines 1 to 55, columns 45 to 105.

The parallelisation of the code of the model has to be undertaken so that the performance of the

model is speeded up. The aim of the model being the short-term forecast of the motion of the clouds, the computational speed is a crucial issue.

The split of images into ‘windows’ has to be investigated further: encouraging results were found despite of the ‘over estimation of brightness’ of the common corners between overlapped ‘windows’. Moreover, working on ‘windows’ is at the basis of the parallelisation of the model.

As an alternative to the parallelisation of the model or as a complement, the variational Bayes approach introduced by Barillec [2] could be integrated to the model.

In the framework of the thesis, the advection field is cell-dependent: each cell has its own movement. It would be interesting to develop cell-independent advection-field defined for the whole image.

## 6.4 Afterword

The project has been interesting for many reasons: from the abstract theory to the practical applications. It has not only be a way to understand and apply the methods learnt in the first term of the MSc, but it has also been a way to discover the underneath processes behind a domain unfortunately taken for granted: meteorology. Moreover, this year has been a precious overview of the three years to come: let us hope they will be at least as interesting and enriching as these nine months of research.

# Appendix A

## Digital to geographic coordinates conversion routine

The equation of the viewing line can be defined in vector terms as:  $OS + k.SI'$  where  $OS$  is the vector from the origin to the spacecraft,  $k$  is a scalar variable and  $SP$  is a vector of any size defining the direction from the spacecraft towards the viewed point. Varying  $k$  then defines the locus of all points on the line.

The vector  $OS$  has coordinates  $(R_s, 0, 0)$ <sup>1</sup>.

The vector  $SP$  can be determined from the line and pixel coordinates of the point in the image. These coordinates are directly related to the angular offset of the viewing line from the equator and the sub-spacecraft meridian.

The vector  $SP$  has coordinates  $(p, q, r)$ .

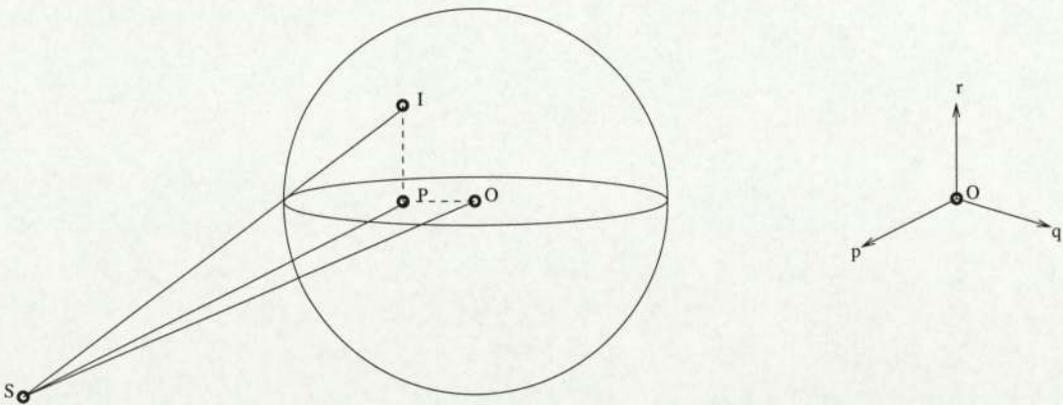


Figure A.1: Viewing line from the spacecraft: one intersection. O is the centre of the Earth, S the satellite, I the observed point and P the projection of I in the Cartesian coordinate system of the Earth.

$(O, p, q, r)$  is the Cartesian coordinates system defined as follows:

---

<sup>1</sup> $R_s$  is the distance from the satellite to the origin.

## APPENDIX A. DIGITAL TO GEOGRAPHIC COORDINATES CONVERSION ROUTINE

- origin at the centre of the Earth;
- $x$  axis in the direction from the origin to the spacecraft;
- $y$  axis to the right as seen from the spacecraft;
- $z$  axis towards the top (north) as seen from the spacecraft.

Let  $\alpha_L$  be the angle  $\widehat{PSI}$  and  $\alpha_P$  be the angle  $\widehat{PSO}$ .

The MATLAB code for the conversion routine is as follows:

```
function [visible, longitude, latitude] = pixl2geo(line, pixel)
% PIXL2GEO used to perform the digital to geographical conversion
%
% visible - flag set to 1 is pixel is on the visible disc
%          - flag set to 0 if pixel is in space
% longitude - longitude of the pixel (degrees East from Greenwich)
% latitude - latitude of the pixel (degrees North from Equator)
%
% line - line number, measured from southern end of frame
% pixel - pixel number measured from eastern pixel of frame

% Set up constants
% altitude - distance from Earth centre to satellite
% req - equatorial Earth radius
% rpol - polar Earth radius
% oblate - Earth oblateness

altitude = 42164.0;
req = 6378.140;
rpol = 6356.755;
oblate = 1.0/298.257;

deg_to_rad = pi/180.0;
rad_to_deg = 180.0/pi;
```

Step is the radiometer step as seen by the spacecraft, in degrees. The image represents an  $18^\circ * 18^\circ$  field of view divided up on an equi-angular basis. For this program an VIS channel of  $400 * 900$ .

```
step = 18/2500;
```

Convert line/pixel values to angular offsets from centre point.

```
aline = ((line + (1996 - 1)) - 1250)*step;
```

```
asamp = (1250 - (pixel + (761 - 1)))*step;
```

```
aline = aline*deg_to_rad;
```

```
asamp = asamp*deg_to_rad;
```

APPENDIX A. DIGITAL TO GEOGRAPHIC COORDINATES CONVERSION ROUTINE

```
% Calculate tangents of angles
```

```
tanal = tan(aline);
```

```
tanas = tan(asamp);
```

Calculate components of an arbitrary vector from the spacecraft in the viewing direction.

```
p = -1;
```

```
q = tanas;
```

```
r = tanal*sqrt(1 + q*q);
```

The location of the point on the Earth can be identified by solving a quadratic equation for the intersection between the Earth's surface and the viewing line from the spacecraft. If this equation has no real roots then there is no intersection; otherwise the required root is the one nearer to the spacecraft (on the visible side of the Earth).

```
a = q*q + (r*req/rpol)^2 + p*p;
```

```
b = 2*altitude*p;
```

```
c = altitude*altitude - req*req;
```

Calculate discriminant. If it is negative (no real roots to quadratic equation) there is no intersection between the line of sight and the disc and so the pixel does not correspond to visible data.

```
det = b*b - 4*a*c;
```

```
if (det < 0)
```

```
    visible = 0;
```

```
    latitude = 0;
```

```
    longitude = 0;
```

```
else
```

```
    visible = 1;
```

```
    k = (-b -sqrt(det))/(2*a);
```

```
    x = altitude + k*p;
```

```
    y = k*q;
```

```
    z = k*r;
```

```
    longitude = atan(y/x);
```

```
    cenlat = atan(z*cos(longitude)/x);
```

```
% This is the geographic latitude. Convert it to the geodetic
```

```
% (or geographic) latitude before returning it to the calling program.
```

```
latitude = atan(tan(cenlat)/(1 - oblate)^2);
```

```
% Convert from radians to degrees.
```

```
latitude = latitude*rad_to_deg;
```

```
% longitude + 10: rectified image 10degree East %
```

```
longitude = longitude*rad_to_deg + 10;
```

```
end
```

## Appendix B

# Maximum Likelihood and EM algorithm

The negative log-likelihood for a general mixture model is:

$$\begin{aligned} E &= -\ln \mathcal{L}(\Theta), \\ &= -\sum_{n=1}^N \ln \left\{ \sum_{j=1}^M p(\mathbf{x}|j, \theta_j) P(j) \right\}. \end{aligned}$$

The individual parameter  $\theta_j$  is identified with each component in the mixture. To find the maximum likelihood estimators for each set of parameters  $\theta_j$ , the error measure is differentiated.  $M$  equations of the form:

$$\frac{\partial E}{\partial \theta_j} = -\sum_{n=1}^N P(j|\mathbf{x}_n) \frac{\partial}{\partial \theta_j} \ln p(\mathbf{x}_n|j),$$

have to be solved in the mixture model case, one for each set of component parameters  $\theta_j$  [17].

Considering the case where the component densities are Gaussian (of spherical covariance  $\Sigma_j = \sigma_j^2 \mathbf{I}$ ):

$$p(\mathbf{x}|j) = \frac{1}{(2\pi\sigma_j^2)^{d/2}} \exp\left\{-\frac{(\mathbf{x} - \mu_j)^2}{2\sigma_j^2}\right\},$$

each of the component densities has two parameters,  $\mu_j$  and  $\sigma_j^2$ . The derivatives of the error measure with respect to  $\theta_j$  at the maximum likelihood solution are given by:

$$\hat{\mu}_j = \frac{\sum_{n=1}^N P(j|\mathbf{x}_n) \mathbf{x}_n}{\sum_{n=1}^N P(j|\mathbf{x}_n)}, \quad (\text{B.1})$$

$$\hat{\sigma}_j^2 = \frac{\sum_{n=1}^N P(j|\mathbf{x}_n) (\mathbf{x}_n - \hat{\mu}_j)^2}{\sum_{n=1}^N P(j|\mathbf{x}_n)}, \quad (\text{B.2})$$

$$\hat{P}(j) = \frac{1}{N} \sum_{n=1}^N P(j|\mathbf{x}_n). \quad (\text{B.3})$$

As the maximum likelihood estimates cannot be found directly from these equations because each parameter appears implicitly in the right-hand side, an iterative re-estimation algorithm will be used to find the maximum likelihood parameters.

**EM algorithm**

The principle of the Expectation-Maximisation algorithm as described in [4] is as follows. Making some initial guess for the parameters of the Gaussian mixture model, called the ‘old’ parameters values, allows the evaluation of the right-hand sides in B.1, B.2 and B.3 which gives a revised estimate for the parameters, called the ‘new’ parameter values, for which the value of the error function is smaller. These parameter values then become the ‘old’ values, and the process is repeated. This reasoning gives the basis for the general two-step EM algorithm:

**E-step:** the expected complete-data log-likelihood with respect to the distributions  $p(\mathbf{z}_n|\mathbf{x}_n, \theta^{old})$ ,  $\mathcal{E}^{comp}(\theta^{new})$ , is computed using the current, fixed, values of the parameters;

**M-step:** the new, re-estimated, values of parameters  $\theta^{new}$  are determined by maximising the value of  $\mathcal{E}^{comp}(\theta^{new})$  computed in the E-step.

In the Gaussian mixture model, the data being considered to be incomplete (the component  $j$  having generated a given data point  $n$  is unknown) a variable  $\mathbf{z}_n$  is introduced. The latter takes on integer values in the range  $1 \dots M$  and denotes the unknown generating component.

Then, the complete-data log-likelihood is given by:

$$\begin{aligned} \mathcal{L}^{comp}(\theta) &= \sum_{n=1}^N \ln p(\mathbf{x}_n, \mathbf{z}_n | \theta), \\ &= \sum_{n=1}^N \ln \{p(\mathbf{x}_n | \mathbf{z}_n, \theta) P(\mathbf{z}_n | \theta)\}. \end{aligned}$$

Taking the expectation of the complete-data log-likelihood with respect to the distribution  $P(\mathbf{z}) = \prod_{n=1}^N P(\mathbf{z}_n | \mathbf{x}_n, \theta^{old})$  and noting that since  $\mathbf{z}$  is a discrete variable, the expectation over all  $\mathbf{z}_n$  is a combination of  $N$  sums:

$$\begin{aligned} \mathcal{E}^{comp}(\theta) &= \sum_{n=1}^N \left[ \sum_{z_1=1}^M \sum_{z_2=1}^M \dots \sum_{z_N=1}^M \prod_{m=1}^N P(\mathbf{z}_m | \mathbf{x}_n, \theta^{old}) \ln \{p(\mathbf{x}_n | \mathbf{z}_n, \theta) P(\mathbf{z}_n | \theta)\} \right], \\ &= \sum_{n=1}^N \sum_{j=1}^M P(j | \mathbf{x}_n, \theta_j^{old}) \ln \{p(\mathbf{x}_n | j, \theta_j^{new}) P^{new}(j)\}, \end{aligned}$$

nothing that  $p(\mathbf{x}) = \sum_{j=1}^M p(\mathbf{x}|j, \theta_j)P(j)$  and that  $P(\mathbf{z}_n, \theta^{new})$  is the prior  $P^{new}(j)$  (the details of the computation can be found in [4] or [17]).

So, after making an initial guess for the parameters, the EM algorithm for a Gaussian mixture model is thus:

**E-step:** computation of the posterior responsibilities  $P(j|\mathbf{x}_n)$  for every data point  $n$  and every mixture component  $j$  from Bayes rule, using the current values of the parameters  $\mu_j^{old}, (\sigma_j^2)^{old}$  and  $P^{old}(j)$ :

$$P(j|\mathbf{x}_n) = \frac{p(\mathbf{x}_n | \mu_j^{old}, (\sigma_j^2)^{old}) P^{old}(j)}{\sum_{j=1}^M p(\mathbf{x}_n | \mu_j^{old}, (\sigma_j^2)^{old}) P^{old}(j)}.$$

**M-step:** computation of the re-estimated values for the parameters  $\mu_j^{new}, (\sigma_j^2)^{new}$  and  $P^{new}(j)$ :

$$\begin{aligned} \mu_j^{new} &= \frac{\sum_{n=1}^N P(j|\mathbf{x}_n, \theta^{old}) \mathbf{x}_n}{\sum_{n=1}^N P(j|\mathbf{x}_n, \theta^{old})}, \\ (\sigma_j^2)^{new} &= \frac{\sum_{n=1}^N P(j|\mathbf{x}_n, \theta^{old}) (\mathbf{x}_n - \mu_j^{new})^2}{\sum_{n=1}^N P(j|\mathbf{x}_n, \theta^{old})}, \\ P^{new}(j) &= \frac{1}{N} \sum_{n=1}^N P(j|\mathbf{x}_n, \theta^{old}). \end{aligned}$$

## Appendix C

# Initialisation of the network: example on a single window

The purpose of this Appendix is to illustrate the functioning of the initialisation algorithm (Section 4.2) on a single weighted ‘window’. The latter (Figure C.1) is the image obtained after having performed the discrimination cloudy/non-cloudy regions and having weighted the overlap common parts shared with its neighbours.

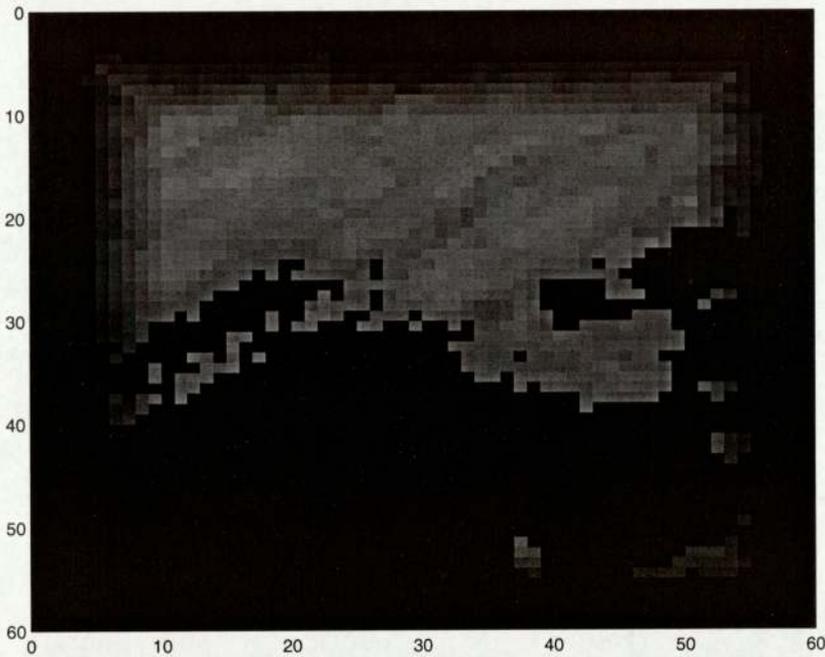


Figure C.1: Original weighted ‘window’.

The first step of the algorithm consists in the localisation of the brightest pixel of the image. The coordinates of the centre of the first radial basis function are fixed to those coordinates. Then, after a local optimisation has been performed, the effect of the first modelled ‘cloud cell’ is removed from  $I_{Temp}$  (Figure C.2). The Figures C.3 and C.4 are the images obtained after the removing of the two and three first ‘cloud cells’.

As soon as the stopping criterion has been reached, no radial basis function is added. The temporary network  $I_{Temp}$  has no more relevant information (Figure C.5).

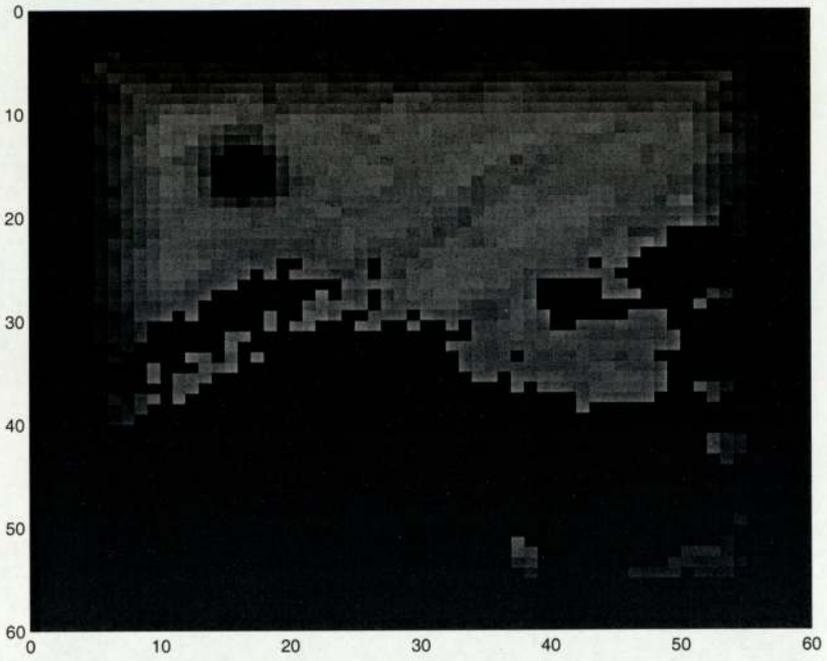


Figure C.2: Removing of the effect of the first radial basis function.

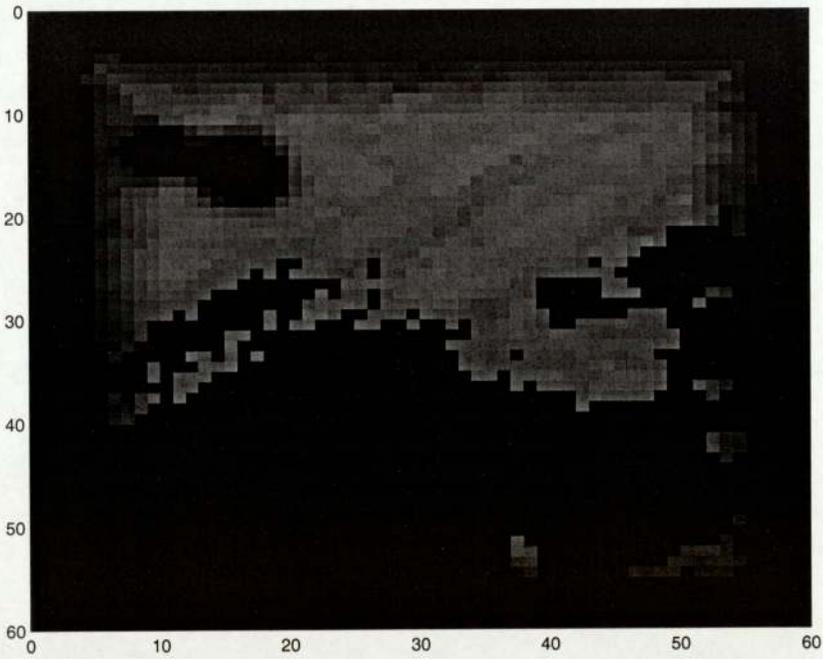


Figure C.3: Removing of the effect of the two first radial basis functions.

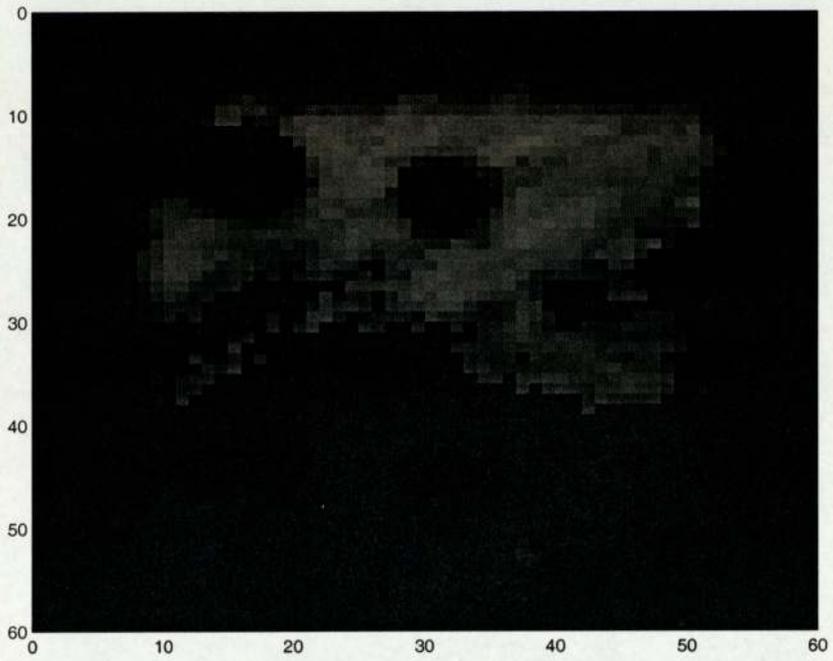


Figure C.4: Removing of the effect of the three first radial basis functions.

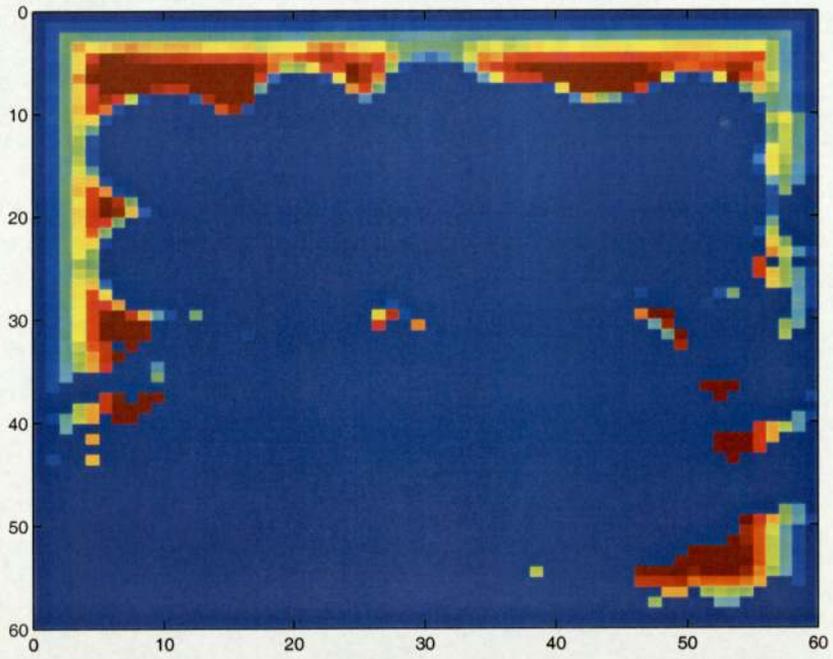


Figure C.5: Stopping criterion reached.

APPENDIX C. INITIALISATION OF THE NETWORK: EXAMPLE ON A SINGLE WINDOW

Now that the steps 1 to 7 of the initialisation algorithm have been undertaken, its last step, the global optimisation of the network, has to be performed. The optimisation of the parameters of the radial basis functions altogether can be modelled by the moving of the centres of the radial basis functions (C.6). The first guess centres are represented by circles whereas the centres after the global optimisation are represented by crosses. The 'moving' of the centres is modelled by a line.

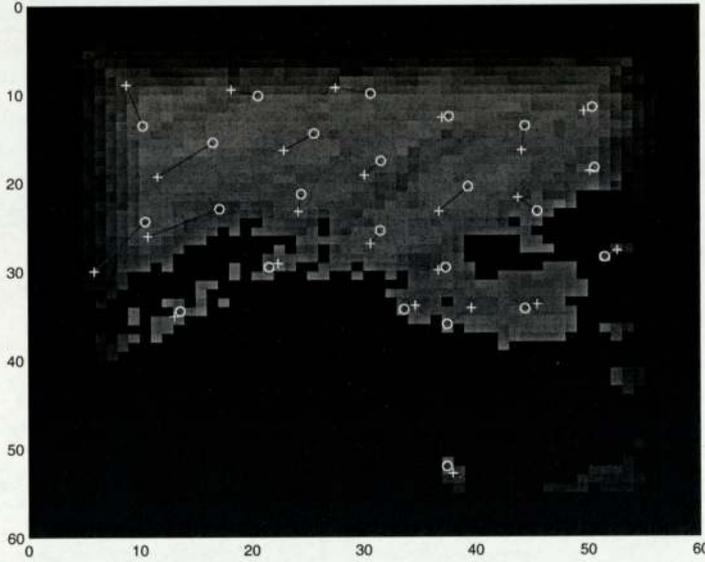


Figure C.6: Moving of the centres after the global optimisation of the network has been performed.

Finally, the modelled cloud is presented in Figure C.7. The crosses are the centres of the modelled 'cloud cells'.

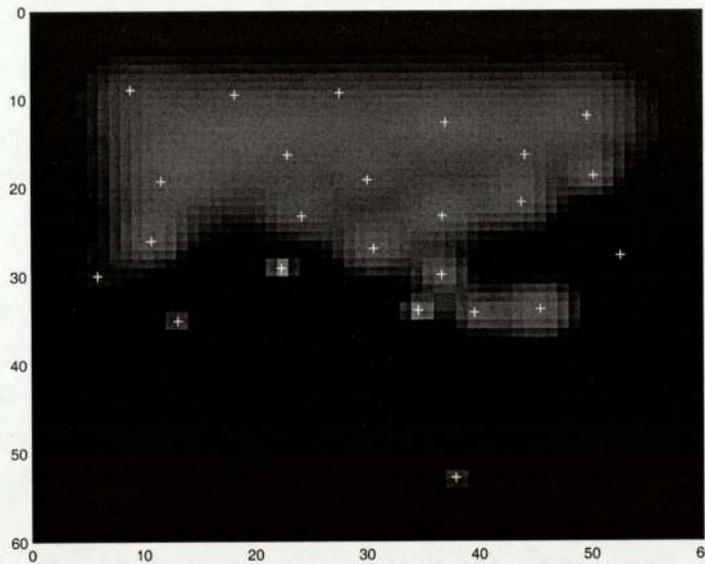


Figure C.7: Modelled cloud and centres of the modelled 'cloud cells'.

# Appendix D

## Kalman filter

The purpose of this appendix is to introduce the Kalman filter for a linear model.

### Measurement or observation equation:

For a discrete linear model, the *measurement* equation is as follows:

$$I_t = F_t \theta_t + \epsilon_t^I,$$

where  $\epsilon_t^I$  is a Gaussian noise term with mean zero and covariance matrix  $\Sigma_t^\epsilon$ .

### State or transition equation:

The state vector can evolve between time steps according to the *state* equation:

$$\theta_{t+1} = E_t \theta_t + W_t \nu_t$$

$\nu_t$  is a white noise term of covariance matrix  $\Sigma_t^\nu$ .

## D.1 Time update: Evolution step

Between two observations, the state evolves according to the *state* equation. Therefore, the estimates of  $\theta$  and  $\Sigma^\theta$  have to evolve while waiting for the new observation. The probability distribution of  $\theta_{t+1}$  given all previous observations up to and including time  $t$ ,  $\mathcal{I}_t$ , is conditioned on  $\theta_t$ :

$$\begin{aligned} p(\theta_{t+1}|\mathcal{I}_t) &= \int d\theta_t p(\theta_{t+1}, \theta_t|\mathcal{I}_t) \\ &= \int d\theta_t p(\theta_{t+1}|\theta_t) p(\theta_t|\mathcal{I}_t) \end{aligned} \tag{D.1}$$

Assuming a Gaussian distribution for  $\theta$  at time  $t$  given  $\mathcal{I}_t$ ,  $p(\theta_t|\mathcal{I}_t) = \mathcal{N}(\hat{\theta}_{t|t}, \hat{\Sigma}_{t|t}^\theta)$  and using the *state* equation:

$$p(\theta_{t+1}|\theta_t) = \mathcal{N}(E_t \theta_t, W_t \Sigma_t^\nu W_t^T).$$

Substituting the latter equation in Equation D.1, the evolution equations for the state and the covariance estimates are:

$$\begin{aligned} \hat{\theta}_{t+1|t} &= E_t \hat{\theta}_{t|t}, \\ \hat{\Sigma}_{t+1|t}^\theta &= E_t \hat{\Sigma}_{t|t}^\theta E_t^T + W_t \Sigma_t^\nu W_t^T. \end{aligned}$$

## D.2 Measurement update: Assimilation step

In the case of a Bayesian approach, updating the model comes down to finding the *maximum a posteriori* estimate of the state at time  $t$  given the new observation  $I_t$  and the estimate of the state at the previous time step (time  $t - 1$ ). Using Bayes' rule, the posterior distribution can be written:

$$p(\theta_t | \mathcal{I}_t) = \frac{p(I_t | \theta_t) p(\theta_t | \mathcal{I}_{t-1})}{p(I_t | \mathcal{I}_{t-1})},$$

where  $\mathcal{I}_t$  denotes all the observations up to and including time  $t$ .

Assuming Gaussian distributions, the densities are defined as follows:

**Likelihood:**  $p(I_t | \theta_t) = \mathcal{N}(F_t \theta_t, \Sigma_t^\epsilon)$ , the observation is centred around the observed process's value with uncertainty due to the noise;

**Prior:**  $p(\theta_t | \mathcal{I}_{t-1}) = \mathcal{N}(\hat{\theta}_{t|t-1}, \hat{\Sigma}_{t|t-1}^\theta)$ , the distribution is centred on the estimate of the state vector at time  $t$  given all the observations up to  $t - 1$ ;

**Evidence:**  $p(I_t | \mathcal{I}_{t-1}) = \mathcal{N}(\hat{I}_{t|t-1}, \hat{\Sigma}_t^I)$ , the mean of the distribution the best prediction of the new observation given the previous observations;

**Posterior :**  $p(\theta_t | \mathcal{I}_t) = \mathcal{N}(\hat{\theta}_{t|t}, \hat{\Sigma}_{t|t}^\theta)$ , the posterior distribution is a Gaussian centred on the best state estimate given all the data up to and including time  $t$ . Both state vector and covariance matrix are updated using the new observation.

In order to find  $\hat{\theta}_{t|t}$  and  $\hat{\Sigma}_{t|t}^\theta$ , the negative log-posterior is minimised with respect to  $\theta_t$ :

$$\begin{aligned} -\log p(\theta_t | \mathcal{I}_t) &\propto (I_t - F_t \theta_t)^T (\Sigma_t^\epsilon)^{-1} (I_t - F_t \theta_t) \\ &+ (\theta_t - \hat{\theta}_{t|t-1})^T (\hat{\Sigma}_{t|t-1}^\theta)^{-1} (\theta_t - \hat{\theta}_{t|t-1}) \\ &- (I_t - \hat{I}_{t|t-1})^T (\Sigma_t^I)^{-1} (I_t - \hat{I}_{t|t-1}), \end{aligned} \quad (\text{D.2})$$

and

$$-\log p(\theta_t | \mathcal{I}_t) \propto (\theta_t - \hat{\theta}_{t|t})^T (\hat{\Sigma}_{t|t}^\theta)^{-1} (\theta_t - \hat{\theta}_{t|t}). \quad (\text{D.3})$$

Equating the terms in  $\theta_t$  in Equations D.2 and D.3 leads to:

$$(\hat{\Sigma}_{t|t}^\theta)^{-1} = F_t^T (\Sigma_t^\epsilon)^{-1} F_t + (\hat{\Sigma}_{t|t-1}^\theta)^{-1}.$$

Differentiating D.2 with respect to  $\theta_t$ , setting to zero (optimum) and then incorporating D.3 to the result gives:

$$\begin{aligned} \hat{\theta}_{t|t} &= (F_t^T (\Sigma_t^\epsilon)^{-1} F_t + (\hat{\Sigma}_{t|t-1}^\theta)^{-1})^{-1} ((\hat{\Sigma}_{t|t-1}^\theta)^{-1} \hat{\theta}_{t|t-1} + F_t^T (\Sigma_t^\epsilon)^{-1} I_t) \\ &= \hat{\theta}_{t|t-1} + K_t \mathbf{e}_{t|t-1}, \end{aligned}$$

where

$$K_t = \hat{\Sigma}_{t|t}^\theta F_t^T (\Sigma_t^\epsilon)^{-1}$$

and

$$\begin{aligned} \mathbf{e}_{t|t-1} &= I_t - F_t \hat{\theta}_{t|t-1} \\ &= \theta_t - \hat{\theta}_{t|t-1} \end{aligned}$$

$K_t$  is called the Kalman gain and  $\mathbf{e}_{t|t-1}$  is the prediction error vector (called the innovation).

# Bibliography

- [1] Asmir Asif. Data assimilation in large time-varying multidimensional fields. *IEEE Transactions on Image Processing*, 8(11), Nov 1999.
- [2] Rémi Barillec. Improving a stochastic rainfall forecasting model. Master's thesis, Aston University, 2003.
- [3] Emmanuel Batail. Statistical models for rainfall: Characterisation and forecasting. Master's thesis, Aston University, 2002.
- [4] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [5] Dan Cornford. A bayesian state space modelling approach to probabilistic quantitative precipitation forecasting. Technical report, Aston University, 2002.
- [6] J. Dozier and S. I. Outclat. An approach toward energy balance simulation over rugged terrain. *Geographical Analysis*, (11):65–85, 1979.
- [7] Geir Evensen. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, (53):343–367, 2003.
- [8] Frédéric Lê. Spatial modelling of cloud. Master's thesis, Aston University, 2003.
- [9] H. G. Lewis, S. Côté, and A. R. Tatnall. Determination of spatial and temporal characteristics as an aid to neural network cloud classification. *International journal of remote sensing*, 18(4):899–915, 1997.
- [10] J. Le Marshall, L. Leslie, R. Morison, N. Pescod, R. Seecamp, and C. Spinoso. Recent developments in the continuous assimilation of satellite wind data for tropical cyclo track forecasting. *Adv. Space Res.*, 25(5):1077–1080, 2000.
- [11] Dipti Prasad Mukherjee and Scott T. Acton. Cloud tracking by scale space classification. *IEEE Transactions on Geoscience and Remote Sensing*, 40(2), Feb 2002.
- [12] G. Ohring, S. Lord, J. Derber, K. Mitchell, and M. Ji. Applications of satellite remote sensing in numerical weather and climate prediction. *Adv. Space Res.*, 30(11):2433–2439, 2002.
- [13] Christophe Papin, Patrick Boutheymy, and Guy Rochard. Unsupervised segmentation of low clouds from infrared METEOSAT images based on a contextual spatio-temporal labeling approach. *IEEE Transactions on Geoscience and Remote Sensing*, 40(1), Jan 2002.
- [14] James J. Simpson. Improved cloud detection and cross-calibration of astr, modis and meris data. Technical report, Digital Image Analysis Laboratory, Scripps Institution of Oceanography, University of California, San Diego, 1999.

## BIBLIOGRAPHY

- [15] James J. Simpson and Jason I. Gobat. Improved cloud detection for daytime avhrr scenes over land. *Remote sensing of environment*, (55):21–49, 1996.
- [16] The METEOSAT Archive. User handbook. Technical report, EUMETSAT, Nov 2001.
- [17] Mike Tipping. Statistical pattern analysis: Lecture notes, Feb 1998.
- [18] Greg Welch and Gary Bishop. An introduction to the kalman filter. Technical report, University of North Carolina, Apr 2004.

# Index

- EUMETSAT, 10, 15
- METEOSAT, 10, 14
  
- advection, 33
  - equation, 33, 35, 38
  - field, 33, 35, 38
- advection field, 33
- albedo, 17, 20, 23
  
- brightness, 13, 17, 19, 20, 24, 26
  - of the ground, 15
  
- cloud albedo, 13
- cloud cells, 25, 26, 28, 30, 33, 34, 38
- cloud field, 33, 38
- cloud optical thickness, 13, 23, 47
- cloud rate, 33, 35, 38
- coordinates, 14, 15
  
- dataset, 12, 17, 19, 37, 47
  
- EM algorithm, 19
- evolution, 35
  
- forecast, 38
- FRS, *see* Full Rapid Scan
- Full Rapid Scan, 10
  
- Gaussian, 17, 36–38
  - basis function, 25
- Gaussian Mixture Model, 17, 19, 20, 24, 47, 48
  
- Kalman filter, 6, 9, 34, 35, 38
  
- land albedo, 13
- Laplace approximation, 36
- latitude, 13, 14, 26
- longitude, 13, 14, 26
  
- measurement update, 34
- minimum brightness, 17
- modelling, 12
- motion of the clouds, 6, 33, 35, 49
  
- Negative log-likelihood, 36
- Negative log-prior, 36
  
- observation equation, 34
- observation model, 20
  
- percentage of cloudiness, 13–15, 20
- priors, 38
  
- Rapid Scanning Service, 10
- RBF, 9, 25, 26
- reflectance, 14
  - of the clouds, 13
  - of the ground, 12–15, 17, 19, 20, 23, 47, 48
- RSS, *see* Rapid Scanning Service
  
- Scaled Conjugate Gradient, 27
- short-term
  - forecasting, 6, 48
  - rainfall forecasting, 6, 48
- solar flux density, 14
- solar irradiance, 13, 14
- space and time model, 13, 37
- space-time model, 25
- static model, 9, 13
  
- threshold, 20, 23
- time update, 34
- transition equation, 34
  
- update, 38
  
- VIS, *see* visible images
- visible images, 10, 12, 13, 23
  
- windows, 27, 39