



Algorithm Theoretical Basis Document  
for the Automatic Satellite Image  
Interpretation Processor of the  
NWC/GEO

**Code:**

NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII

**Issue:** 2.0 **Date:** 21 January 2019

**File:** NWC-CDOP2-GEO-ZAMG-SCI-ATBD-  
ASII\_v2.0.docx

**Page:**

1/40

The EUMETSAT  
Network of  
Satellite Application  
Facilities



**NWC SAF**

Support to Nowcasting and  
Very Short Range Forecasting

# Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO

NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII, Issue 2, Rev. 0

21 January 2019



*Applicable to*

*GEO-ASII-v2.4.2 (NWC-046b)*

**Prepared by ZAMG**



### REPORT SIGNATURE TABLE

Function	Name	Signature	Date
<b>Prepared by</b>	A. Wirth, V.Zwatz-Meise (ZAMG)		<i>21 January 2019</i>
<b>Reviewed by</b>	A.Jann, V.Meyer (ZAMG)		<i>21 January 2019</i>
<b>Authorised by</b>	Pilar Rípodas (AEMET) <i>SAFNWC Project Manager</i>		<i>21 January 2019</i>

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 3/40
---	---	---	---

## DOCUMENT CHANGE RECORD

Version	Date	Pages	CHANGE(S)
2.2	23 February 2009	41	Initial version (content derived from various precursor documents on PGE10, most notably “Software User Manual for the SAFNWC/MSG: Scientific part for the PGE10”)
2.3	22 September 2009	43	Update for release v2010
2.4	2 February 2011	44	Update for release v2011
2.4.1	15 February 2012	42	No scientific updates. Date issue and rev. adapted to v2012
2.4.2	15 July 2013	41	Update for release SAFNWC/MSG v2013, after thorough independent review of document logic, comprehensibility etc.
2.0 (CDOP-3)	21 January 2019	40	Update for release SAFNWC/MSG v2018 taking into account CDOP-3 nomenclature and similar formal aspects

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 4/40
---	---	---	---

## Table of contents



<b>1. INTRODUCTION .....</b>	<b>7</b>
1.1 SCOPE OF THE DOCUMENT .....	7
1.2 SOFTWARE VERSION IDENTIFICATION .....	7
1.3 IMPROVEMENT FROM PREVIOUS VERSION .....	7
1.4 DEFINITIONS, ACRONYMS AND ABBREVIATIONS .....	7
1.5 REFERENCES .....	8
1.5.1 <i>Applicable Documents</i> .....	8
1.5.2 <i>Reference Documents</i> .....	9
<b>2. AUTOMATIC SATELLITE IMAGE INTERPRETATION (ASII) OVERVIEW .....</b>	<b>10</b>
<b>3. AUTOMATIC SATELLITE IMAGE INTERPRETATION (ASII) ALGORITHM DESCRIPTION .....</b>	<b>11</b>
3.1 THEORETICAL DESCRIPTION .....	11
3.1.1 <i>Satellite data processing</i> .....	11
3.1.1.1 Pre-processing of imagery .....	11
3.1.1.2 Production and evaluation of Atmospheric Motion Vectors (AMV) .....	13
3.1.1.3 Derivation of local statistical parameters .....	15
3.1.2 <i>Development images</i> .....	15
3.1.2.1 Algorithm .....	15
3.1.2.2 Practical applications .....	15
3.1.2.3 Application to conceptual models .....	15
3.1.3 <i>Detection of contiguous areas with similar cloud structure</i> .....	16
3.1.3.1 Algorithm for image segmentation .....	16
3.1.3.2 Classification of the derived image segments .....	17
3.1.3.3 Application to Conceptual Models .....	19
3.1.4 <i>NWP data input</i> .....	20
3.1.5 <i>Utilization of analysis information from the preceding slot ("memory")</i> .....	21
3.1.5.1 Algorithm .....	21
3.1.5.2 Practical applications .....	21
3.1.6 <i>Attribute fields</i> .....	21
3.1.6.1 Algorithm .....	21
3.1.6.2 Practical applications .....	22
3.1.7 <i>Neighbourhood-related functions</i> .....	23
3.1.7.1 Neighbourhood check .....	23
3.1.7.2 Removal of "not sufficiently widespread" signals .....	23
3.1.7.3 Removal of non-compact structures .....	23
3.1.7.4 Gap filler .....	23
3.1.7.5 Hole filler .....	24
3.1.7.6 Displacement .....	24
3.1.7.7 Assessment of a nearby value .....	25
3.1.7.8 Extension function .....	25
3.1.8 <i>Identification of coherent areas: drawing of contour lines and determination of frontal areas</i> .....	26
3.1.8.1 Algorithms .....	26
3.1.8.2 Practical applications .....	27
3.1.8.3 Practical usage in ASII .....	27
3.1.9 <i>Frontal rear side detection</i> .....	27
3.1.9.1 Algorithm .....	27
3.1.9.2 Practical usage in ASII .....	28
3.1.10 <i>Identification of S-shaped lines</i> .....	28
3.1.10.1 Algorithm .....	28
3.1.10.2 Practical applications .....	29
3.1.11 <i>Identification of concave and convex sections of a contour line</i> .....	29
3.1.11.1 Algorithm .....	29
3.1.11.2 Practical applications .....	30



<b>3.1.12</b>	<b><i>Detection of convective cells</i></b> .....	<b>30</b>
3.1.12.1	Algorithm .....	30
3.1.12.2	Practical applications .....	30
<b>3.1.13</b>	<b><i>Dark Stripe detection</i></b> .....	<b>31</b>
3.1.13.1	Algorithm .....	31
3.1.13.2	Practical applications .....	31
<b>3.1.14</b>	<b><i>Detection of fibres</i></b> .....	<b>32</b>
3.1.14.1	Algorithm .....	32
3.1.14.2	Practical applications .....	32
<b>3.1.15</b>	<b><i>Evaluation of texture and spiral structures for CM classification</i></b> .....	<b>33</b>
3.1.15.1	Algorithms .....	33
3.1.15.2	Practical applications .....	33
3.1.15.3	Applications to conceptual models.....	34
<b>3.2</b>	<b>PRACTICAL CONSIDERATIONS</b> .....	<b>34</b>
<b>3.2.1</b>	<b><i>Validation</i></b> .....	<b>34</b>
<b>3.2.2</b>	<b><i>Output format description</i></b> .....	<b>35</b>
3.2.2.1	The conceptual model output .....	35
3.2.2.2	The AMV output .....	37
<b>4.</b>	<b>ASSUMPTIONS AND LIMITATIONS</b> .....	<b>39</b>
<b>5.</b>	<b>REFERENCES</b> .....	<b>40</b>

## List of Tables and Figures

<b>Table 1:</b> List of Applicable Documents .....	9
<b>Table 2:</b> List of Referenced Documents .....	9
<b>Table 3:</b> Description of classes. The column “number” gives code numbers which are occasionally referred to in later sections. ....	18
<b>Table 4:</b> List of conceptual models using the image classification .....	19
<b>Table 5:</b> ASII BUFR Output Products Names.....	36
<b>Table 6:</b> AMV Output Products Names.....	37
<b>Figure 1:</b> An example of an ASIINWP product to indicate the common grid spacing for the automatic satellite image interpretation module. The full resolution is best assessed from the ‘ul’ (=“upper level low”) indicators over Scandinavia and the dot symbols around Albania (other CM labellings were partly thinned to improve the readability). ....	15
<b>Figure 2:</b> Illustration of the classification set-up. First diagram: In a first training step, local statistical parameters were stratified by subjectively assigned categories (distinguished in the diagram by different colours; the numbers in the legend refer to the classes defined in <b>Table 3</b> ). Second diagram: Actual classification: The average statistical parameters of segments are compared with those average values of each class that were derived from the first diagram. The final classification is shown here. ....	19
<b>Figure 3:</b> Illustration of the principle of transforming the information of a detected maximum into a probability field. The blue isolines represent the field of the smoothed WV-AMV-vorticity, the red concentric lines represent the thresholds of the derived probability field. ....	22
<b>Figure 4:</b> IR-satellite image from the 22nd October 2001, 12 UTC. The green line encircles the rear side of the CF. The batch of red lines mark the region enlarged in the direction of the orthogonal wind (see below) covering most parts of the cold front cloud band. The lines are generated by displacement. The protrusion of the CF at the Pyrenees is caused by lee cloudiness not physically separated from the frontal cloud band and therefore added to the latter. ....	25
<b>Figure 5:</b> Schematic indicating the features being relevant for the S-line detection.....	29
<b>Figure 6:</b> Configuration investigated during the test for black stripes. The pixel value at point X is compared with values at 8 selected pixels( $A_r$ to $H_r$ ) at several search radii $r$ . ....	32
<b>Figure 7:</b> Schematic of a frontal system with texture elements (black) and Hough lines (orange). ....	33

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 7/40
---	---	---	---

## 1. INTRODUCTION

### 1.1 SCOPE OF THE DOCUMENT

This document is the Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation (ASII) which is the PGE10 of the SAFNWC/MSG software package.

This document contains a scientific description of the PGE10 algorithms, including the needed input data and scientific aspects.

### 1.2 SOFTWARE VERSION IDENTIFICATION

This document describes the algorithms implemented in the PGE10 version contained in the 2018 SAFNWC software package delivery.

### 1.3 IMPROVEMENT FROM PREVIOUS VERSION

No changes in the PGE as such, which is scientifically frozen. However, the new version allows NWP input from the GFS model (in addition to ECMWF input which has been supported from the very beginning).

### 1.4 DEFINITIONS, ACRONYMS AND ABBREVIATIONS

AD	Applicable Document
AMV	Atmospheric Motion Vector
ASII	Automatic Satellite Image Interpretation
ASIINWP	Automatic Satellite Image Interpretation including NWP fields
BUFR	Binary Universal Form from the Representation of meteorological data
CA	Cold air Advection
CAC	Cold Air Cloudiness
Cb	Cumulonimbus
CC	Convective Cells
CEOF	Complex Empirical Orthogonal Function
CF	Cold Front
CM	Conceptual Model
COST	European Cooperation on Scientific and Technical Research
DI	Dry Intrusion
EC	Enhanced Cumulus
ECMWF	European Centre for Medium-range Weather Forecasts
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FI by Jet	Front intensification by Jet crossing
GFS	Global Forecast System, the weather forecast model of the National Centers for Environmental Prediction

IR	Infrared
JI	Front Intensification by Jet Streak
MSG	Meteosat Second Generation
NWP	Numerical Weather Prediction
OCCL	Occlusion
PGE	Product Generation Element
PV	Potential Vorticity
PVA	Positive Vorticity Advection
RC	Rapid Cyclogenesis
RSV	Relative Shear Vorticity
RV	Relative Vorticity
SAF	Satellite Application Facility
SAFNWC	SAF on support to NoWCasting and Very-Short-Range Forecasting
SAI	Stability Analysis Imagery
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SH	Showalter Index
SUM	Software User Manual
TA	Temperature Advection
TFP	Thermal Front Parameter
ULL	Upper Level Low
UM	User Manual
UTC	Universal Time Coordinated
WA	Warm air Advection
WF	Warm Front
WV	Water Vapour
ZAMG	Zentralanstalt für Meteorologie und Geodynamik



## 1.5 REFERENCES

### 1.5.1 Applicable Documents

For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the document referred to applies.

Latest documentation can be found at the SAFNWC Help Desk at <http://nwc-saf.eumetsat.int>



 	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 9/40
---	---	--

Ref	Title	Code	Vers	Date
[AD.1]	Project Plan for the NWCSAF CDOP3 phase	NWC/CDOP3/SAF/AEMET/MGT/PP	1.0	06/03/18
[AD.2]	NWCSAF CDOP3 Project Plan Master Schedule	NWC/CDOP3/SAF/AEMET/MGT/PP/Ma sterSchedule	1.1	28/02/18
[AD.3]	Configuration Management Plan for the NWC SAF	NWC/CDOP3/SAF/AEMET/MGT/CMP	1.0	21/02/18
[AD.4]	System and Components Requirements Document for the NWC/GEO	NWC/CDOP2/GEO/AEMET/SW/SCRD	2.1	21/01/19
[AD.5]	Interface Control Document for Internal and External Interfaces of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/ICD/1	1.0	21/01/19
[AD.6]	Interface Control Document for the NWCLIB of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/ICD/2	1.0	21/01/19
[AD.7]	Data Output Format	NWC/CDOP3/GEO/AEMET/SW/DOF	1.0	21/01/19
[AD.8]	Component Design Document for the NWCLIB of the NWC/GEO	NWC/CDOP2/GEO/AEMET/SW/ACDD/ NWCLIB	2.0	27/02/17
[AD.9]	NWC SAF Product Requirements Document	NWC/CDOP3/GEO/AEMET/MGT/PRD	1.0	01/18
[AD.10]	User Manual for the Tools of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SCI/UM/To ols	1.0	21/01/19

*Table 1: List of Applicable Documents*



## 1.5.2 Reference Documents

The reference documents contain useful information related to the subject of the project. These reference documents complement the applicable documents. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the document referred to applies.

Latest documentation can be found at the SAFNWC Help Desk at <http://nwc-saf.eumetsat.int>.

Reference	Title	Code	Vers	Date
[RD.1]	Validation Report for the “Automatic Satellite Image Interpretation” (ASII-PGE10, v2.0)	SAF/NWC/CDOP/ZAMG/SCI/VR/ 01	1.1	28/01/08
[RD.2]	Scientific Report on changes implemented in “Automatic Satellite Image Interpretation” (ASII- PGE10 v2.1)	SAF/NWC/CDOP/ZAMG/SCI/RP/ 01	1.0	16/11/07
[RD.3]	Validation Report for “Automatic Satellite Image Interpretation” (ASII-PGE10, v2010)	SAF/NWC/CDOP/ZAMG/SCI/VR/ 4	1.0	20 April 2010
[RD.4]	User Manual for the Automatic Satellite Image Interpretation Processor of the NWC/GEO: Science Part	NWC/CDOP3/GEO/ZAMG/SCI/UM/A SII	1.0	21/01/19

*Table 2: List of Referenced Documents*

 	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 10/40
---	---	--

## 2. AUTOMATIC SATELLITE IMAGE INTERPRETATION (ASII) OVERVIEW

PGE10 of the SAFNWC/MSG provides an automatic satellite image interpretation in terms of conceptual models (CMs). Conceptual models are important tools for forecasters in diagnosis, nowcasting and very short range forecasting; they describe a synthesis of typical features in satellite images as well as in other synoptic material (for instance numerical model output parameters) and the physical processes causing them. A CM diagnosis cannot only be used for detection of an ongoing process but also for the determination of the stage of development within this process.

An inventory of the usage of CMs in weather services was produced within the COST78 project, providing a collection of more than 80 CMs used in European weather services as well as a collection of the existing experiences with them (COST78, 1996). Other literature providing summaries of CMs include Bader et al. (1995) and ZAMG (1999, 2009).



Typical features/configurations in satellite images are fundamental input to conceptual models and often show the very first stages of newly developing cloud features. To be able to recognise all necessary satellite features and feature changes within a time appropriate for nowcasting, the only way to manage the huge amount of satellite data (images) is an automatic detection of CMs.

PGE10 carries out the recognition of CMs in two ways (two products):

- ASII: Automatic Satellite Image Interpretation from MSG SEVIRI satellite data alone
- ASIINWP: Automatic Satellite Image Interpretation from MSG SEVIRI satellite data supplemented by useful key parameters from the numerical model output (Numerical model used at ZAMG: ECMWF; however, the NWCSAF application allows to use another model, GFS).

The meteorologist can classify and analyse a satellite image due to his knowledge and expertise. He recognises the position and shape of frontal systems, the brightness of image pixels and the extension of synoptic systems. The implemented methods and algorithms aim to reflect all the elements (criteria) contained in a CM, i.e. its characteristic pattern, brightness and texture, its location compared to other features in the satellite image, the input from NWP fields and more. The combination of these input parameters results in a selection of image features and assigns the analysed object to a certain category.

This document describes the implemented methods and algorithms which are fundamental for the detection of the chosen CMs. How these methods are combined to yield the final output is described in the Product User Manual for this PGE (see [RD.4]).

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 11/40
---	---	---	--

## 3. AUTOMATIC SATELLITE IMAGE INTERPRETATION (ASII) ALGORITHM DESCRIPTION<sup>1</sup>

### 3.1 THEORETICAL DESCRIPTION

#### 3.1.1 Satellite data processing

##### 3.1.1.1 Pre-processing of imagery

Satellite imagery need to be pre-processed before they are passed on to the pattern recognition procedures of the ZAMG SAFNWC software package. The necessary input data comprise the MSG spectral channels WV6.2, IR10.8 and IR12.0. Necessary steps are:

- Conversion of the radiance into the meteorologically more meaningful quantity radiation temperature with help of the calibration factors of the sensor(s). This is implemented by calling the respective NWCLIB functions (cf. [AD.8]).
- Correction for differences in brightness caused by differing geographical location. IR images show the purely latitudinal effect that similar systems tend to appear brighter in Northern Europe than over the Mediterranean (because of the over-all colder atmosphere there). For certain components of the pattern recognition software - which deals with all parts of Europe - this has to be taken into account by darkening the northern area in a suitable fashion. A simple empirical formula was derived on the basis of manual classifications where patterns considered as representing the same phenomenon were collected for a large number of cases. The derived regression relations of the basic parameters in dependence on latitude and season were included into the software to provide correction coefficients for each individual pixel.
- The shape of cloud systems tends to be different between the North and the South on the original satellite projection due to purely geometric effects associated with the spherical shape of the earth. The employed remedy is to transform the images to a cylindrical projection.

##### 3.1.1.1.1 Image filter

###### ◆ Binomial Smoothing Filter

Let  $p_{ij}$  denote the input image of dimension  $i \times j$ . Our linear filters for image processing return a smoothed image  $S$ , with the smoothing operation defined in an  $m \times n$  matrix  $M$ :



$$S(p_{ij}) = M * p_{ij},$$

(\* designates the two-dimensional convolution operation). Hence, at any pixel, the resulting value of the filter is a weighted sum of the surrounding  $m \times n$  pixels.

Currently, 3x3, 5x5, 7x7, and 9x9 binomial filters are in use (depending on the application). The advantage of these two-dimensional filters is their separability, i.e. they can be efficiently implemented as the subsequent computation of a row and a column filter, respectively:

---

<sup>1</sup> Several of the algorithms described in the manual have been developed in the frame of scientific projects supported by the Austrian Ministry for Education, Science and Culture in cooperation with the company GEPARD.

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 12/40
---	---	---	--

$$M = B_3 = \vec{b}_3^T * \vec{b}_3 = \frac{1}{4} \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} * \frac{1}{4} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

for the 3×3 filter; the larger filters are analogously constructed, with:

$$\vec{b}_5 = \frac{1}{16} \begin{pmatrix} 1 & 4 & 6 & 4 & 1 \end{pmatrix}$$

$$\vec{b}_7 = \frac{1}{64} \begin{pmatrix} 1 & 6 & 15 & 20 & 15 & 6 & 1 \end{pmatrix}$$

$$\vec{b}_9 = \frac{1}{256} \begin{pmatrix} 1 & 8 & 28 & 56 & 70 & 56 & 28 & 8 & 1 \end{pmatrix}$$

These binomial filters are used for

- Elimination of spurious pixels
- Increasing the robustness of gradient search methods

#### ◆ Sobel filter

The Sobel operator is a non-linear edge filter, which describes the local variability of the field of grey shades.

Let  $p_{ij}$  denote the input image of dimension  $i \times j$ . First, two new (“partial-derivative”) images are constructed via (the symbol  $*$  again designates the two-dimensional convolution operation):

$$S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} * p_{ij}, S_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} * p_{ij}.$$



From these, the final Sobel image is obtained by

$$S(p_{ij}) = \sqrt{S_x^2 + S_y^2}.$$

The Sobel operator is used to describe the local texture of patterns by visualizing the magnitude of the intensity gradients seen in IR satellite imagery. It is one of the two inputs for the segmentation process, described in more detail in section 3.1.1.3, where areas of similar optical appearance are grouped. Furthermore, the Sobel operator is a component of the recognition algorithm for fibres and the enhancement of edges within the contouring algorithms. For the latter application, the Sobel-filtered image is linearly combined with the original image. This is done in order to enhance the contrast along cloud band edges.

#### 3.1.1.1.2 Gaussian pyramid

With the Gaussian pyramid a sequence of images of successively decreasing size is obtained by the recursive application of the steps described hereafter. It is a useful method to transform large-scale images into a sequence of successively smaller scaled images in order to allow the

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 13/40
---	---	---	--

application of algorithms derived for small-scale objects (high resolution necessary) to large-scale objects in the reduced images. Starting with the original image  $I_0$ :

(a) The production of the supplementary image  $H_{n+1}$  through smoothing the image of the  $n$ -th step  $I_n$  with a suitable  $5 \times 5$ -filter, e.g.  $B_5$ .

(b) The production of the  $(n+1)$ -th image  $I_{n+1}$  by taking only every second pixel of every second line from the supplementary image  $H_{n+1}$ . Thus, the size of the image  $I_{n+1}$  is a quarter of the size of the image of previous step  $I_n$ . The synoptic information content (shape of synoptic cloud systems) is still well recognisable, even with compression factors as high as 16 or 64.

This method ranks among the standards for image processing to diminish objects under the side condition of minimum loss of information (see Jähne (1993)).

Gaussian pyramids allow

- to apply the same detection algorithms on image objects of quite different spatial extension, and
- to drastically reduce computation time by the diminution of the image size.

### **3.1.1.2 Production and evaluation of Atmospheric Motion Vectors (AMV)**

The ASII software employs the standard cross-correlation technique applied to two temporally consecutive images. By minimizing cross-correlation functions, rectangular subareas of the temporally preceding image (so-called targets) are linked to the most similar subarea of the subsequent image (matcher). The method follows closely that described in Kidder and Vonder Haar, 1995, pp. 236ff). The basic materials are pairs of cylindrically projected IR or WV images. The projection has the advantageous geometrical effect that the rectangular targets can be constructed even far up in the north. This allows deriving AMVs even at high latitudes. In literature, reservations can be found (Hodges, 1998) regarding the application of cross-correlation methods to large systems (over large time steps). However, these reservations do not apply here, for though the ASII product requires vectors representing the movement of larger areas, the considered spatial resolution is high enough that the displacement of a whole synoptic system is determined.

In the ASII package, the AMVs are computed on a regular grid for every 16-th pixel (~110 km). The tracked target regions are rectangular subareas taken from the first image with a side length of 33 pixels. The AMV computations in PGE10 follow a three-step scheme (a similar approach has also been advocated by Szantai et al., 2000):

1. Derivation of first-guess vectors on images being twice reduced in size by the Gaussian pyramid algorithm, which offers the advantage of drastically reduced size but with almost identical information content (cf. section 3.1.1.1.2; size of the image: 1/16 the size of the original image). Target size is  $9 \times 9$  pixels. Search region is  $9 \times 9$  pixels, which corresponds to an area of  $36 \times 36$  pixels of the original image (i.e. maximum displacements of 18 pixels in positive or negative  $x$ - or  $y$ -direction).
2. Derivation of a modified vector on images being reduced in size by a single application of the Gaussian pyramid algorithm (size of the image: 1/4 of the original). A region of  $5 \times 5$  pixels around the first-guess displacement is considered as potential AMV end points. Target size is  $17 \times 17$  pixels of the reduced image.
3. Derivation of the ultimate vector on the original images, as described above.

The displacement with the best match sometimes gives a spurious vector (e.g. because patterns change distinctively by cloud dissolution or formation and “displacement” is essentially the wrong concept then). It was found that in the ZAMG implementation, such outliers appear rather seldom



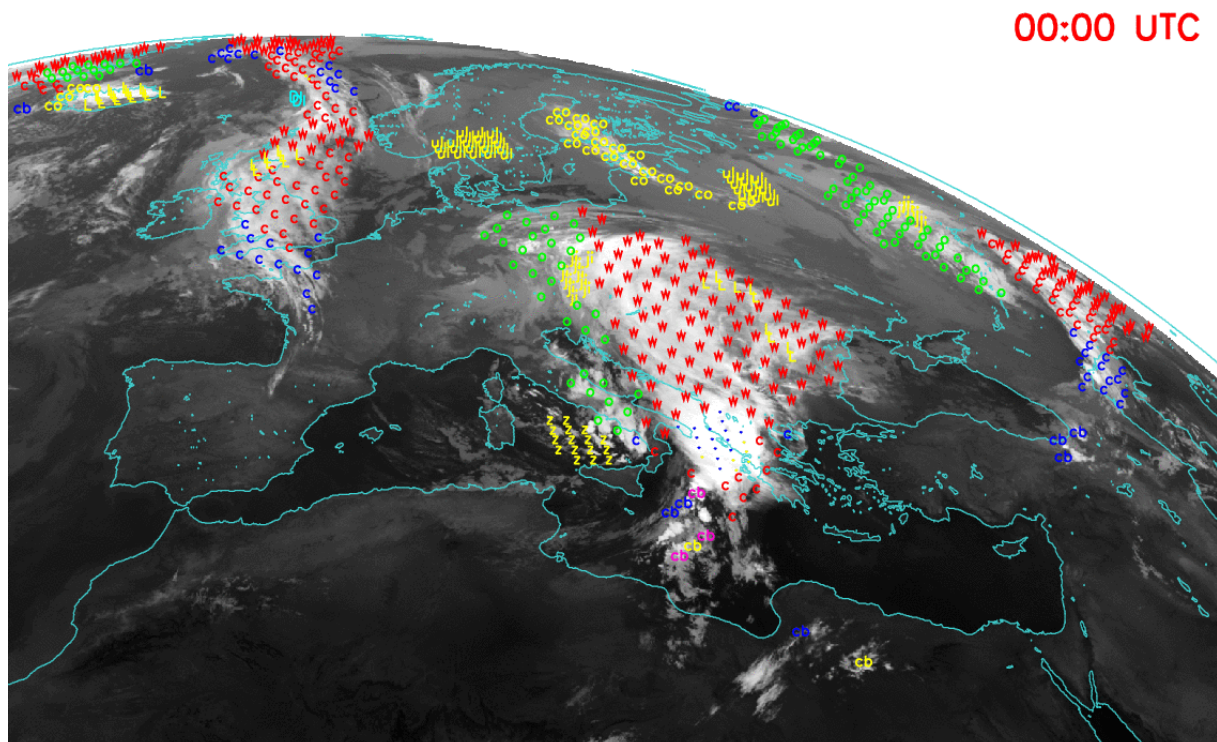
and can be effectively filtered. The usage of a regular grid on which the 'best matches' are calculated allows to keep the filter very simple. Currently the u- and v- component of the vectors undergo a repeated (threefold) filtering with a 3x3-median filter.



The AMV fields calculated from the used IR and WV channel are optionally available as separate output. To activate this option, the respective entries have to be properly set in the model configuration file for PGE10 (see the ICD/1 document for the External and Internal Interfaces of the SAFNWC/MSG and also 3.2.2.2). Unlike other AMV products, this one provides a field of smoothed vectors appropriate for nowcasting purposes. Through considerable post-processing (smoothing operations described in the following), single spurious vectors are eliminated and subsequent quality checks (e.g. based on spatial coherence) are not necessary. Height information of AMVs has not been required for those nowcasting applications that are known so far (extrapolation of surface observations, image pixels etc.) and is therefore omitted.

The AMVs are provided on the standard analysis grid (**Figure 1**). From the WV-AMV-fields, rotation (i.e. vorticity) and divergence are computed. A raw and a smoothed version are available for each parameter (the raw version corresponds to small-scale phenomena whereas the smoothed parameters correspond to the synoptic-scale flow). The pattern recognition within PGE10 relies on patterns like distinct maxima or the mere sign of a quantity. Therefore, no effort was undertaken to achieve conformity to the SI units of vorticity and divergence. Rather, the computations base on simple approximations like  $\Delta v_x - \Delta u_y$  (for vorticity), where  $\Delta v_x = v(x+\Delta x, y) - v(x-\Delta x, y)$  and  $\Delta u_y = u(x, y+\Delta y) - u(x, y-\Delta y)$ . All figures stated in the following refer to the arbitrary units resulting from this proceeding.

The „smoothed“ vorticity and divergence fields are obtained by applying a 7x7 pixel-wide boxcar average filter to both the u- and the v-component of the AMVs before the finite differences are calculated as described before.

The AMVs themselves are often important ingredients for neighbourhood-related functions (described below) where the vectors are used on several occasions to indicate the search direction.



		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 15/40
---	---	---	--

*Figure 1: An example of an ASIINWP product to indicate the common grid spacing for the automatic satellite image interpretation module. The full resolution is best assessed from the 'ul' (= "upper level low") indicators over Scandinavia and the dot symbols around Albania (other CM labellings were partly thinned to improve the readability).*

### **3.1.1.3 Derivation of local statistical parameters**

The software of PGE10 delivers conceptual model classifications on a regular grid in a cylindrical projection covering the entire region defined by the user on the SEVIRI disk. Around each grid point, a rectangle of 12x10 pixels is considered and evaluated to obtain selected local statistical parameters. They constitute the parameters for the subsequent area segmentation (section 3.1.3.1). But they are also used directly for pattern recognition purposes. As in all other algorithms described hereafter, precautions are taken to deal with possible image-free regions at the edges properly.

The following parameters are computed for each grid point:

- Mean value
- Maximum value
- Minimum value
- Sobel operator value as a measure of spatial variability (cf. 3.1.1.1.1)
- Standard deviation.

The actual list varies for the input images IR, WV or development image (section 3.1.2.1).

## **3.1.2 Development images**

### **3.1.2.1 Algorithm**



With the help of the AMV field (chapter 3.1.1.2), a SEVIRI image is extrapolated to the next time step (15 or 30 minutes). The development image is calculated as the differences between the extrapolated image and the image of the subsequent time step for each pixel. If only advective changes took place the extrapolated and the actually measured image would be identical and the development image consistently zero. The development image can therefore be interpreted as the non-advective changes in the cloud field. The development image can be used as an additional image channel to which all operations can be applied in the same manner as to the other imagery.

### **3.1.2.2 Practical applications**

The development images are used for CMs which potentially show some stronger developments in particular phases of their life cycle. Depending on the values of the development image, the analysed conceptual model is flagged as developing or decaying.

### **3.1.2.3 Application to conceptual models**

Development images are used in the detection process of frontal waves and cumulonimbus cloudiness (see [RD.4]).

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 16/40
---	---	---	--

### 3.1.3 Detection of contiguous areas with similar cloud structure

#### 3.1.3.1 Algorithm for image segmentation

The procedure of image segmentation divides the infra-red satellite image into sub-regions which are coherent in terms of selected statistical quantities. Adjacent areas in a satellite image have a higher likelihood of belonging together in terms of a meteorological classification when they have

- a) similar brightness (e.g. radiation temperature) and
- b) similar texture.

The first step in the segmentation is the computation of local statistical parameters describing brightness (temperature) and texture of pixels inside predefined rectangles around the individual grid points. The current version of the SAF software has the following adjustments:

- initially as many segments as grid points are generated, each segment consisting of the 12x10-pixels rectangle around the respective grid point.
- Following the computation of the basic statistical parameters computed for a segment (cf. 3.1.1.3), the segmentation module uses the quantity  $\mu+2\sigma$  ( $\mu$ : mean,  $\sigma$ : standard deviation) for brightness (temperature) and the value of the Sobel operator smoothed by a 3x3-median filter as texture parameter.

The reasoning behind the choice is:

- $\mu+2\sigma$ : When using the mean  $\mu$  alone, segments occur which comprise only cloud field boundaries with medium values of this parameter. It could eventually happen in the further course of the ASII decision tree that such separated segments are erroneously diagnosed as a CM not found anywhere else among the neighbours. In order to remedy this, we exploit the fact that the character of a border region is clearly reflected in the scatter of the pixel brightness.



Since there is:

- a) large  $\mu$ , but only moderate  $\sigma$  in clouds,
- b) medium  $\mu$  and significant  $\sigma$  in areas of cloud system borders,
- c) small  $\mu$  and low  $\sigma$  in cloudfree areas,

the distinction between group a) and b) can be overcome when an empirical amount of the scatter is added to the mean. Using a parameter like  $\mu+2\sigma$ , transition zones are merged with the adjacent cloud systems.

- Median-smoothed Sobel-filtered image The Sobel operator identifies edges by identifying brightness gradients: it exhibits small values in areas of smooth cloud appearance, larger values when the cloudiness is structured, and very large values at well-defined edges such as boundary regions of cloud shields. There is the danger that these very large values of the Sobel operator lead to a separated segment of cloud boundary areas which has no meaning in terms of conceptual models. As these areas are generally narrow, the problem can be overcome by applying a 3x3-median filter, replacing a very large Sobel value by a more representative medium value from the adjacent "pure cloud" area. Since the resolution of the satellite imagery decreases when going north, there is a tendency for satellite images to appear smoother in higher latitudes. Hence, an empirical correction formula is applied in order to achieve better comparability between segments at different latitudes in Europe.



		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 17/40
---	---	---	--

After the computation of the basic statistical features on the analysis grid, 'region growing' is applied, which means that each segment is compared with its neighbours. Those two neighbours being most similar according to the objective similarity measure  $D$  (see below) are joined to a new and larger segment. The distance  $D$  in feature space is based on the differences of the normalised statistical parameters  $P$  (normalisation accomplished by dividing by the empirically observed scatter of the respective parameter values), and it is computed by

$$D_{A,B} = \max_{i=1,\dots,n} |P_{i,segment A} - P_{i,segment B}|.$$

The number of parameters  $n$  is 2 in our case and the formula reflects the idea that two segments are considered different when they differ significantly in *either* brightness (temperature) *or* texture; hence the maximum absolute value appears in the given formula. The segmentation process continues by rejoining similar adjacent segments until all adjacent segments become too distinctive to be merged any further. An empirical threshold is applied to terminate the segmentation process, so that the following two conditions are reasonably well satisfied:

- 1) a conceptual model should be covered by as few segments as possible (ideally only one)
- 2) a segment must not comprise areas covered by two conceptual models.

Fulfilment of these conditions guarantees that the automatic algorithm identifies the entity of a weather phenomenon in a manner quite similar to the human observer.

The average values of the two governing parameters ( $\mu+2\sigma$ , Sobel) are determined for each segment. They form the basis for the next step - the classification.

### **3.1.3.2 Classification of the derived image segments**

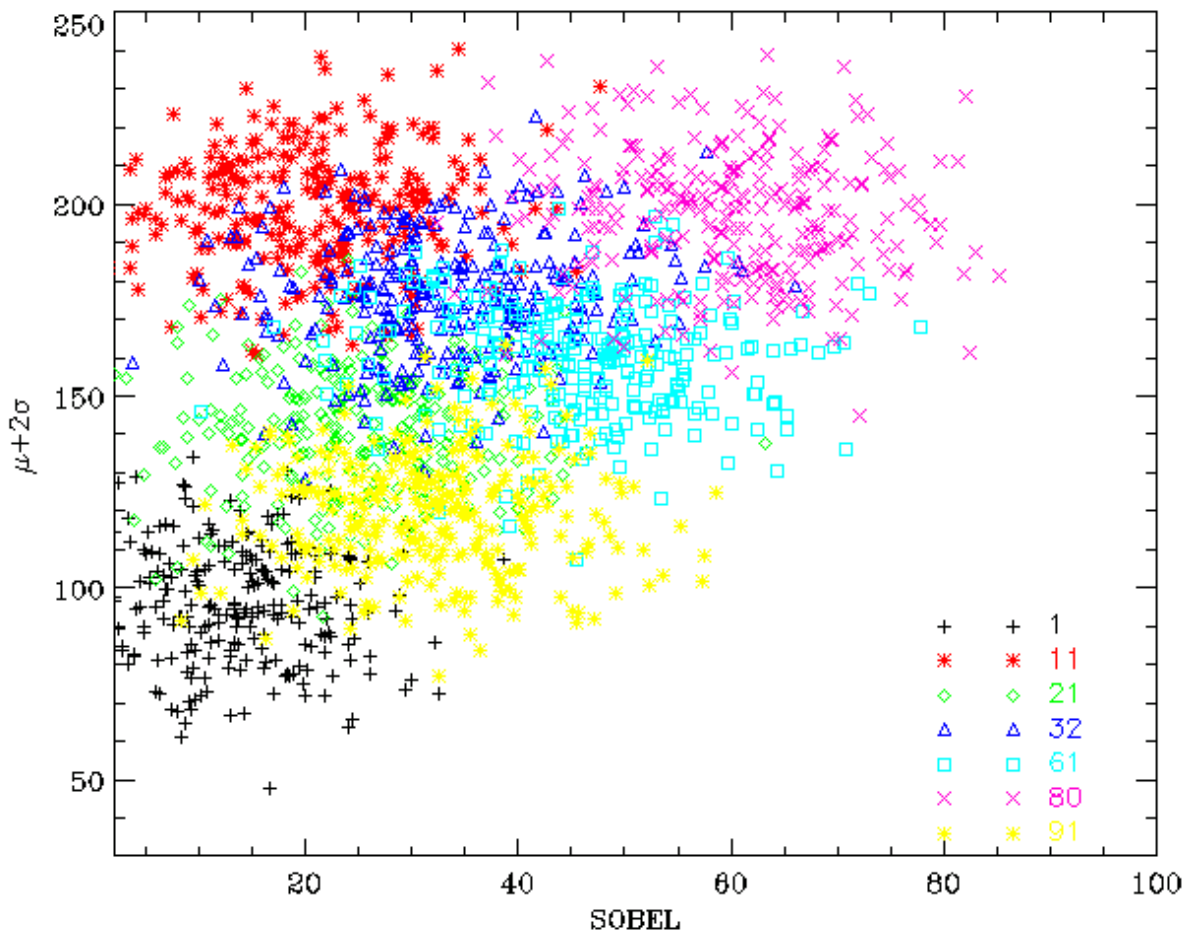
Based on a representative reference sample of manually classified conceptual models, the so-called "centres of gravity", i.e. the average values of the two statistical image features  $\mu+2\sigma$  and the value derived from the application of the median-smoothed Sobel image, were computed for seven selected CM-related classes (**Table 3** lists the chosen classes; Figure 2, diagram 1 shows a schematic). Some of the centres of gravity undergo seasonal variations (not shown), which have been taken into account by empirical corrections. These corrections are constructed as functions of the day of the year. Within the automatic classification process, the most appropriate of the seven centres of gravity is identified for each segment, leading to the designations of the raw classification (Figure 2, diagram 2).

Unfortunately, the patterns in satellite imagery do not form well-separated clusters in the two-dimensional parameter space. Due to the ambiguous classification, the automatic classifications bear an inherent uncertainty. For an improvement by subsequent selective corrections, two additional concepts are available:

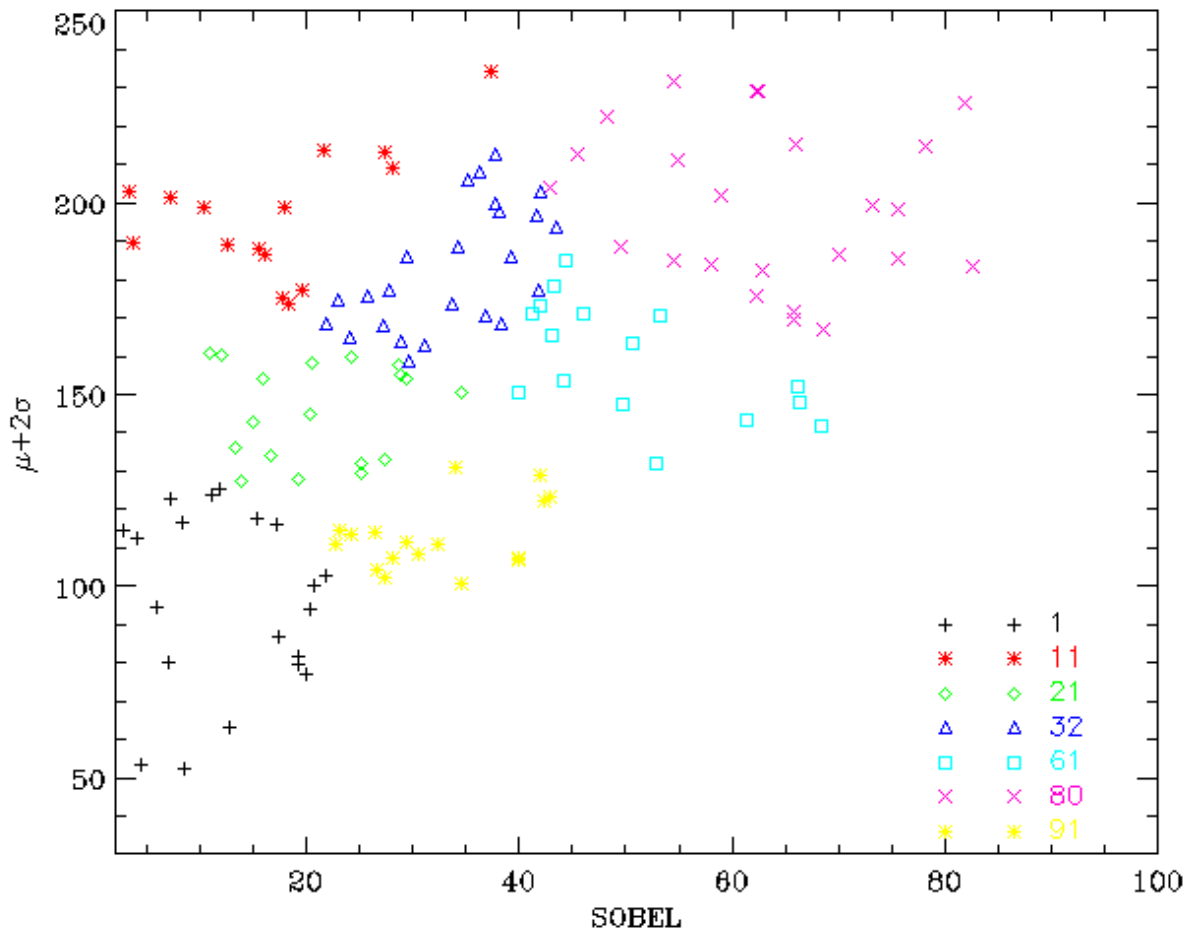
- The second-best classification is taken into account. To find this second-best classification the respective ratio of distances from the two nearest centres of gravity in pattern space is calculated. A ratio close to 1 means that the second-best guess has approximately the same likelihood as the first choice.
- The classification of adjacent segments is considered.

Class	Number
Cloudless area	1
Bright (cold) smooth cloudiness, as characteristic for warm fronts	11
Dark grey to grey cloudiness with smooth to medium rough texture, as it is typically observed in low cold fronts	21
Bright to medium bright cloudiness with rough texture ; as it is typical for multi-layered frontal systems	32
Bright (cold) and small cellular cloudiness as it is typical for enhanced Cumuli	61
Very bright (cold) and bigger cloud cells as they are typical for Cumulonimbus cloudiness	80
Low-level cellular cloudiness as it is typical for the cold-air cellular cloudiness	91

*Table 3: Description of classes. The column “number” gives code numbers which are occasionally referred to in later sections.*



*Figure 2 (first diagram; see legend on following page)*



**Figure 2:** Illustration of the classification set-up. First diagram: In a first training step, local statistical parameters were stratified by subjectively assigned categories (distinguished in the diagram by different colours; the numbers in the legend refer to the classes defined in **Table 3**).  
Second diagram: Actual classification: The average statistical parameters of segments are compared with those average values of each class that were derived from the first diagram. The final classification is shown here.



### 3.1.3.3 Application to Conceptual Models

The classification of image features is used for some of the conceptual models dealt with in ASII. Brightness (temperature) and image texture give a first hint for the type of conceptual model. (The final assignment is done in combination with other parameters, depending on the CM.)

Conceptual Model	Classification
(Low) cold front	21 & 32b*
Mesoscale Convective System (MCS)	11 & 80 in adjacent areas
Cold air cloudiness	61+80b or 32+80b or 80
FI by Jet	11 or 32 or 80

**Table 4:** List of conceptual models using the image classification

\* "b" stands for the second best classification

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 20/40
---	---	---	--

### 3.1.4 NWP data input

At the forecast times 00, 06, 12, and 18 UTC (and - if available - 03, 09, 15 and 21 UTC), the pre-processing of NWP data consists of

- computation of derived quantities;
- interpolation to the analysis grid.



The basic input data of the ASII software package are forecast fields on a regular longitude-latitude grid (during the development phase a 1.5°x1.5°-grid was used); used levels are 1000, 850, 700, 500, 400, 300 hPa. The data are interpolated to the analysis grid by a bilinear scheme. The computation of derived quantities from the basic parameters wind, temperature, geopotential height, and humidity is done within the SAF software. The list of currently used quantities computed at the chosen levels (cf. relevant chapters below and in the corresponding Product User Manual [RD.4]) comprise:

- Geopotential height
- Wind speed
- Temperature advection
- Thermal front parameter
- Vorticity
- Shear vorticity
- Vorticity advection
- Showalter index
- Potential vorticity

For slots with no corresponding NWP forecast, two ways of obtaining the required NWP data exist:

- 1) The “classical” method used in PGE10 since 1996 works via extrapolation using the AMV field calculated for the WV channel (Backward tracking: At each grid point, it is calculated at which point the respective AMV trajectory has been in the preceding image. As this origin normally lies between the the analysis grid prints, the NWP parameters of the preceding slot at the trajectory's origin have to be obtained from spatial interpolation; a bicubic spline interpolation is employed for that purpose).
- 2) An interpolation technique based on Complex Empirical Orthogonal Function (CEOF) analysis can optionally be exploited. Complex Empirical Orthogonal Functions are used to decompose time and space dependent numerical fields into their Eigenmodes. Each Eigenmode consists of a complex vector pair and explains a certain amount of the variability of the field in space and time. A matrix with  $N$  points in space (measurements or grid points) and  $M$  points in time yields  $M$  independent, orthogonal complex vector pairs, comprising a temporal (1) and a spatial function (2):

$$T(t) = A_i(t) \exp[i\Theta(t)] \quad (1)$$

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 21/40
---	---	---	--

$$S(x) = A_s(x) \exp[i\Phi(x)] \quad (2)$$

where  $A_t$  ( $A_s$ ) is the amplitude of the temporal (spatial) wave and  $\Theta$  ( $\phi$ ) its phase. Equation (1) and (2) represent the Eigenmodes of the  $N \times M$  matrix. The most significant modes (i.e. the modes with the highest Eigenvalues) are used to reconstruct the field. The PGE10 implementation requires 4 subsequent forecast outputs at regular time intervals for this interpolation method;  $A_t$  and the phase  $\Theta$  are the parameters which are interpolated in time. The CEOF method is well-suited for the temporal interpolation of model fields, especially when sinusoidal pattern prevail in the data. In these cases, this method clearly shows better results than linear interpolation (e.g. Wirth et al., 2008).

### 3.1.5 Utilization of analysis information from the preceding slot ("memory")

#### 3.1.5.1 Algorithm

The identification of conceptual models often relies on prescribed thresholds. Due to the fixed thresholds, identified CMs occasionally disappear or change into another CM for one time step. This has negative impact on the continuity of analyses. Therefore, a concept has been introduced that memorizes an analysis for exactly one time step. The software dumps the determined information on conceptual models which is then read during the analysis of the following image. The analysis data are extrapolated using the AMV field calculated for the IR channel\* (consult the UM [RD.4] to see which CMs use such a memory function).

#### 3.1.5.2 Practical applications

Two different methods to implement the “memory” algorithm in the automatic detection of a conceptual model are available. The extrapolated ‘old’ and the current diagnosis are compared. Depending on the conceptual model, the process leads to:

1. the retention of the ‘old’ diagnosis
2. a second run of the identification module of the conceptual model, which was indicated in the preceding time step. But this time, less stringent thresholds are used.

### 3.1.6 Attribute fields

#### 3.1.6.1 Algorithm

Some input quantities – originating either from the satellite or from NWP - are incorporated on the basis of attribute fields. When an attribute (e.g. extremum of a field, left exit region) is detected, probability-type fields are computed around the point of interest within a prescribed region of influence. The resulting field exhibits values from 100% at the location of the detected phenomenon from where it gradually decreases towards 0% with an adjustable slope.

---

\*Backward tracking: For each grid point, it is calculated at which point the respective AMV trajectory has its origin in the preceding image. As this origin normally lies in-between the points of the analysis grid, the parameters of the preceding slot at the trajectory's origin have to be obtained from spatial interpolation; a bicubic spline interpolation is employed for that purpose. At present, the concept only supports Boolean variables. If the interpolated value is above 0.5, the result is set to 1; otherwise, the final result is set to 0 at the affected grid point.

Thresholding functions within ASII / ASIINWP can then be used to convert the spatially smooth probability function into a yes-/no-decision for any grid point.

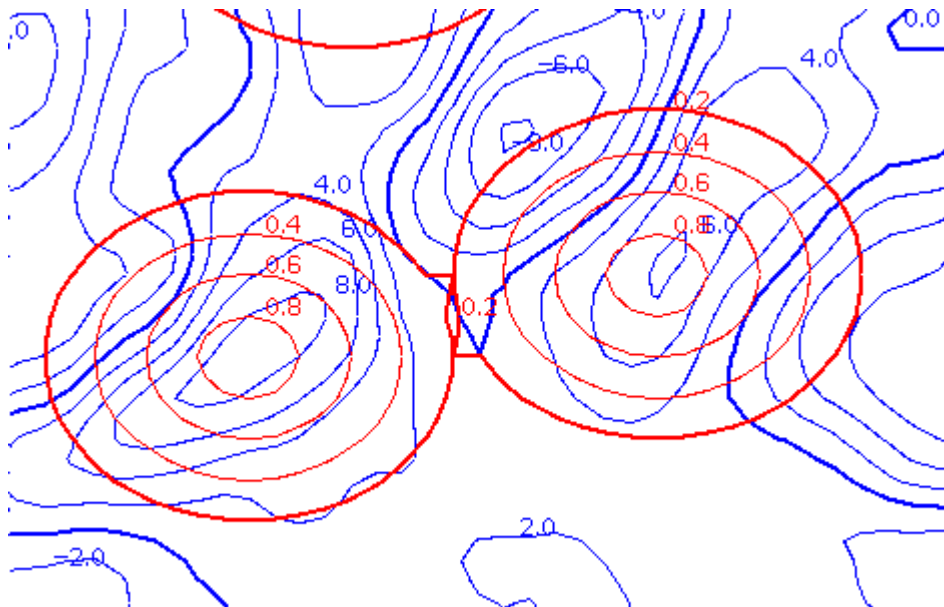
The basic attribute functions which are currently implemented cover the following features:

- local extrema of a field,
- left exit region of a jet streak,
- ridge lines of a field,
- zero lines.

### 3.1.6.2 Practical applications



#### 3.1.6.2.1 Probability fields around a local extremum

Figure 3 illustrates the principle of transforming the information of a detected maximum into a probability field. Maxima are detected in a given numerical field and to each maximum a probability field (the values of which range from 0 to 1) is attributed. In case of overlapping, the higher value is taken. This probability field indicates how near a grid point is located to a maximum. A high value like 0.8 indicates a close maximum, a low value like 0.2 indicates that the respective grid point is far from a maximum. The following parameters can be chosen individually for each application: 1) range and shape of the probability field, 2) minimum magnitude an occurring maximum has to have in order to be taken into account.



**Figure 3:** Illustration of the principle of transforming the information of a detected maximum into a probability field. The blue isolines represent the field of the smoothed WV-AMV-vorticity, the red concentric lines represent the thresholds of the derived probability field.



		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 23/40
---	---	---	---

### 3.1.6.2.2 Probability fields along a line

If the searched pattern is not a single point (e.g. a maximum) but, for example, a zero line, the computation of the probability field is carried out for each grid point which has been identified as a point of the line. The joint probability field comprises the highest achievable probability values for the individual grid points.

### 3.1.6.2.3 Probability regions in relation to jet streaks

The exit region is located in the region where the air flow decelerates at the cyclonic side of the jet streak. Therefore, the zero line of shear vorticity is tracked from the point of maximum wind speed in wind direction until it reaches the location where the wind speed falls below a given threshold. The resulting length of the phenomenon is, together with some adjustable empirical coefficients, input to formulas which fix a point at the cyclonic side of the jet as the centre of the exit region. The attribute field of probabilities around this point is constructed in similarity to 3.1.6.2.1.

## 3.1.7 Neighbourhood-related functions

This group of functions was designed to detect spatial relations between specified sets of ASII grid points. Many of these functions serve as filters, eliminating preliminary diagnoses lying on the wrong side of a certain phenomenon, deleting spurious classifications with too small extension or to fill areas of small extensions in a reasonable manner.

### 3.1.7.1 Neighbourhood check

We consider two fields of Boolean variables (e.g. front area yes/no) X and Y. At each grid point with X=1, the 3x3-surrounding is checked for the presence of Y=1. Only if a sufficient number of suitable points are found, the (Boolean) result variable for the grid point is set to 1.

### 3.1.7.2 Removal of "not sufficiently widespread" signals

In areas of Boolean variable X = 1, which are smaller in size than a given threshold, the variable is set to 0. The threshold is dependent on parameter and CM.

### 3.1.7.3 Removal of non-compact structures



For each grid point with Boolean variable X=1, the 3x3-surrounding is checked for the presence of grid points with the same value. If the number of such neighbours is below a specified threshold, the value for X at the considered grid point is set to 0.

### 3.1.7.4 Gap filler

Starting at a grid point where a Boolean variable has the value 1, another such point is sought in the eight main directions (north, south, east, west, south-east, south-west, north-east and north-west) up to a certain distance. If such a point is found and the points lying on the search path in-between the two points satisfy a certain eligibility criterion, the Boolean variables at all points along the search path are set to 1.

The "gap filler" function is applied when the area of a CM is disrupted by another CM. The filling of such intrusions with the surrounding CM is synoptically reasonable. It is used in the diagnosis process of the following CMs:

- frontal wave
- warm front
- cold front, cold front under warm air advection

 	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 24/40
---	---	---

- occlusion

### 3.1.7.5 Hole filler

Classifications of type X are eliminated and replaced by class Y if:

- a) the area of X is smaller than a specified threshold AND
- b) the area of X is surrounded by points
  - b1) EITHER classified as Y (at least one such point must be present),
  - b2) OR being non-classified.

If these criteria are fulfilled, the large coherent area of classification Y is interrupted by a small region of X, which forms the „hole“ that is to be filled.

In the course of the image analysis, grid points within a CM can incorrectly remain without CM classification, because the underlying data do not satisfy any CM classification criterion. One of the most common phenomena of this kind are holes in a warm front shield caused by CA intrusions at 700 hPa. The hole-filler function adds the missing warm front tags in the hole, provided that the size of the hole does not exceed a certain threshold. This is a synoptically reasonable procedure, provided the hole is not too large and is surrounded by identical tags or if the unanalysed area is located at the margin of a frontal area. Also used in the diagnosis process of cold fronts.

### 3.1.7.6 Displacement

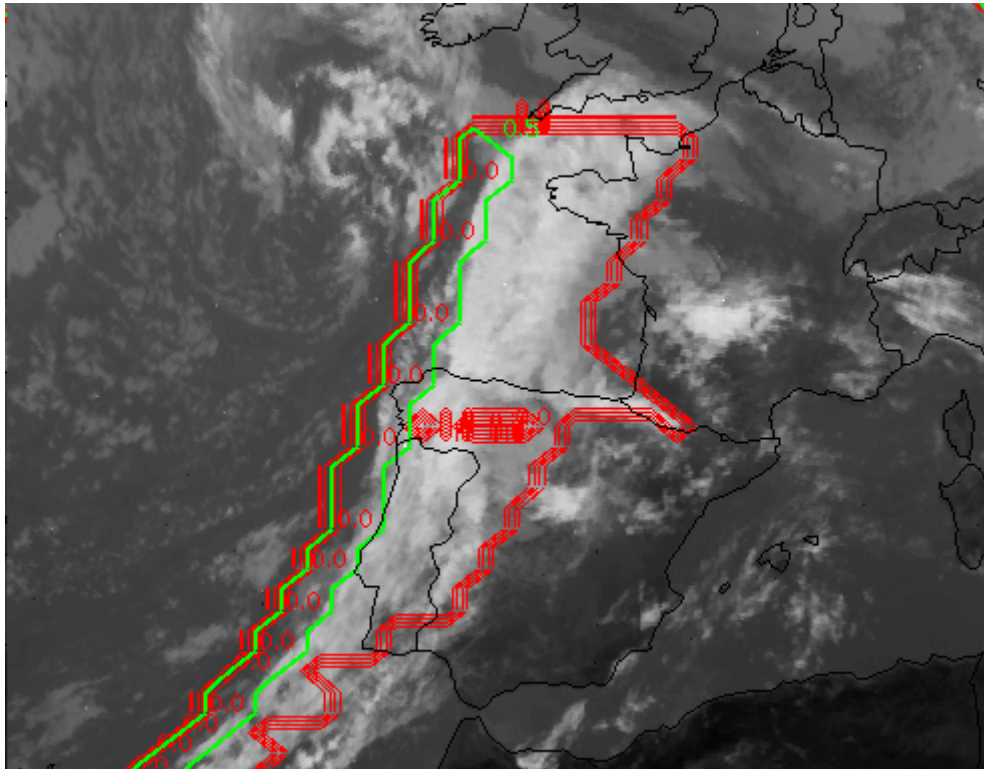
Here, an area where a Boolean variable-of-interest exhibits values of 1 is extended in a certain direction given by an input vector field. The vector field can be uniform over the whole area or vary from one grid point to the next (e.g. when displacing with an AMV field).

The displacement function is used to distribute information, for instance the information of a rearward frontal edge (see Figure 4) downstream for a certain number of grid points. The information “rearward edge” is an indicator for a cold front. Thus, extended to the adjacent cloud band, a cold front can be identified in its entirety.

The “displacement” function is used during the detection of the following CMs:

- lee cloudiness
- frontal wave
- upper wave
- occlusion
- dry intrusion





**Figure 4:** IR-satellite image from the 22nd October 2001, 12 UTC. The green line encircles the rear side of the CF. The batch of red lines mark the region enlarged in the direction of the orthogonal wind (see below) covering most parts of the cold front cloud band. The lines are generated by displacement. The protrusion of the CF at the Pyrenees is caused by lee cloudiness not physically separated from the frontal cloud band and therefore added to the latter.

### **3.1.7.7 Assessment of a nearby value**

In some situations, it is not the neighbouring grid point which is of interest but a distant one. Therefore, a function has been implemented that for every grid point assesses the value of a specific variable in a certain distance and direction and writes the respective results into the output variable at the starting points. The vector field specifying the search direction can be uniform over the whole area or vary from one grid point to the next (e.g. when tracking with an AMV field).



So far, this function is used only for the lee cloud detection to determine if high pixel values on the one side of the mountains are connected with lower pixel values on the other side of the mountains.

### **3.1.7.8 Extension function**

If there is a sufficient number of grid points with a certain classification  $x$  within a given area, which is specified by another variable  $y$ , the classification  $x$  is extended over the whole area. There are two kinds of that functional type: a) extension of a classification over the entire segment where this classification appears; b) extension of a classification over an area of a Boolean  $y$  being 1.

The extension function is used to extend the property of a grid point to adjacent grid points satisfying a certain condition. This function is of particular interest if a key parameter covers only parts of the area of the CM.

The extension function plays a role in the ASII detection of the following CMs:

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 26/40
---	---	---	--

- frontal wave
- occlusion
- dry intrusion
- fibres
- comma cloud
- upper level low

### 3.1.8 Identification of coherent areas: drawing of contour lines and determination of frontal areas

#### 3.1.8.1 Algorithms

The identification of coherent areas of sufficiently large size plays an important role in e.g. the recognition of frontal systems and of S-shaped bulges (waves). Two different - yet related - approaches are followed:

- Contour lines:



The implemented steps to determine a contour line are:

- (1) Scale transformation of the IR image (accomplished by selection of a suitable step of the Gaussian pyramid)
- (2) Enhancement of edges through the combination of the original image with an image obtained by forming its derivatives (cf. Sobel filter in section 3.1.1.1.1).
- (3) Contraction through a recursive median filter
- (4) Drawing of a contour around pixel values above a given threshold in the filtered image
- (5) Thinning of the boundary to obtain a 1-pixel broad contour line
- (6) Elimination of contour lines being too short by using a threshold, that means a contour line has to exceed a minimum length.

- Determination of coherent areas:

This approach identifies the number of contiguous pixels above a certain brightness (temperature) threshold as coherent area. Areas which are too small are eliminated (threshold is depending on CM and was empirically derived). For retained areas, there is the option to add small darker areas (below a prescribed threshold of greyscale values) to the identified bright (cold) area, if fully enclosed by the latter (useful to fill accidental gaps within an identified frontal area).

The identification of coherent areas (e.g. frontal cloud fields) in high latitudes proved difficult during winter times because of cold cloud free land surfaces showing temperatures typical for mid to high level cloudiness. An additional algorithm was therefore implemented: The basic idea is that clouds are always moving at least a little bit and the land surface is stationary. So the differences between 2 consecutive images should be comparatively small in cloudfree and larger in cloudy areas. The best results were obtained by using a threshold for the temperature-difference for channel IR10.8  $\mu\text{m}$  of  $\pm 0.3$  K in the interval of 15 minutes. If the temporal difference is smaller than that threshold for a contiguous area of a minimum extension of 3.000 pixels, then this area is likely to be bare ground, otherwise the bright area is classified as cloud.

 	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 27/40
---	---	---

### **3.1.8.2 Practical applications**

Frontal areas are characterised by their synoptic scale. The size of coherent areas exceeding a certain empirically chosen brightness (temperature) threshold is determined, and the criterion of a minimum size is imposed so that only large cloud systems with a high probability of being frontal are selected. All grid points with at least one assigned pixel passing the test are marked as frontal. The frontal area is constituted by the collection of all these frontal grid points.

The algorithm is applied to the IR image after the application of an image filter and the black stripe detection algorithm. The image filter enhances fibre-like structures by smoothing the image using pixels in the local texture direction (see also 3.1.15). This enhances the contrast between the elongated structures of frontal cloudiness and adjacent areas with less pronounced structure (low homogeneous cloud decks, for example). The black stripe algorithm (cf. 3.1.13) is used to subtract dry areas from frontal regions.

The application of the cloud free land detection in high latitudes during winter seasons has resulted in a better detection of frontal cloud bands in that area. Special attention has to be paid to the periods with enhanced warming/cooling of the surface (after sunrise and near sunset). In these cases misclassification may occur because the change of the surface temperature exceeds the applied thresholds. A future release might see a refined version with temporary asymmetrical thresholds depending on sun zenith angle (note, however, that the ASII PGE is scientifically frozen in the CDOP-2 phase 2012-2017).

Stationary Stratocumulus sheets are also filtered by the algorithm.

### **3.1.8.3 Practical usage in ASII**



The contour lines are used for all conceptual models with frontal characteristics, i.e. all kinds of cold fronts, warm fronts and occlusions. They are also used to locate substructures of cold fronts like waves, dry intrusions, FI by jet or upper waves and to exclude some conceptual models from frontal areas like EC or comma cloudiness.

## **3.1.9 Frontal rear side detection**

### **3.1.9.1 Algorithm**

Many cloud phenomena are characterised by their position with respect to the frontal cloud band. Waves, upper waves and dry intrusions develop only at the rear side of cold fronts. Therefore it is of main importance to distinguish between rear and front side of a frontal cloud system first. This is done with help of the WV AMVs. The algorithm differentiates between frontal and non-frontal regions, indicated by 0 and 1, respectively. Based on this differentiation an advection term is calculated. Described in a simplified manner, the property "non-frontal" is advected with an appropriate component of the WV AMVs and it is checked if the non-frontal area is advected into a frontal area.

The above-described way to detect frontal rear sides has shown to be beneficial for the separation of cold fronts from warm fronts if only satellite image information is available. Investigations have shown that cold fronts are usually characterised by a frontal rear side whilst the latter is missing at warm fronts. In a similar way, wave bulges within cold fronts show a detected rear side at the cold front part of the wave and no detected rear side at the warm front part. Formally spoken, in the case of cold fronts the property "non-frontal" is advected into the frontal area at the cold front part of the wave ('wind' direction taken from WV AMVs) while the property "frontal" is advected from the frontal into non-frontal areas at the WF bulge of the wave.

 	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 28/40
---	---	--

With this separation the problem of finding a synoptically reasonable separation between cold and warm front can be approached. The rear side detection and the width of the frontal area are coupled in order to accomplish such a separation.

A frontal rear side is indicated if the following two conditions are fulfilled:

1. a component of the WV AMV points towards the leading edge of the frontal cloud band;
2. the distance from the detected rear side to the leading edge of the cloud band does not exceed 12 grid points.

### **3.1.9.2 Practical usage in ASII**

Beside the discrimination between warm fronts and cold fronts, the concept of the frontal rear side detection is also applied to other CMs. The frontal rear side is also a criterion for the detection of occlusion cloud bands (cf. the corresponding Product User Manual [RD.4]), since they always have a frontal rear side. This fact is used to make a better differentiation between occlusions and warm fronts.

The criterion of the frontal rear side detection is also used for the localization of CMs existing only at frontal rear sides: the frontal wave, the upper wave, the dry intrusion.

### **3.1.10 Identification of S-shaped lines**

#### **3.1.10.1 Algorithm**

The S-shape of a contour line can be an indication for vigorous meteorological activity, particularly the formation of waves.

The recognition of S shapes is based on the extraction of contours as described in 3.1.8. The algorithm for the identification of S-lines follows the steps:

- (1) Derivation of the contours;
- (2) Through each pixel P lying on a contour line, a batch of straight lines with given radius and angle increments is constructed (see Figure 5);
- (3) It is then investigated, if:
  - one of the straight lines cuts the contour line (after leaving the latter) once in each direction (necessary condition for S-shape),
  - the cutting points, e.g. the point where a straight line cuts the contour line exceed an empirically set minimal distance from the investigated point P, measured along the considered straight line (elimination of tiny „random“ bulges of the contour),
  - the cutting points exceed a certain maximum distance from P, measured along the contour line,
  - one of the two sections of the straight line, which are separated by P, passes through the brighter/colder area, the other through the darker/warmer area,
  - the curvatures of the two contour line sections between P and the two cutting points exceed a threshold which is large enough to justify the designation of an „S“.

(4) Each point with one straight line satisfying the above criteria is marked as centre of an „S“, the associated section of the contour is indicated as „the S“.

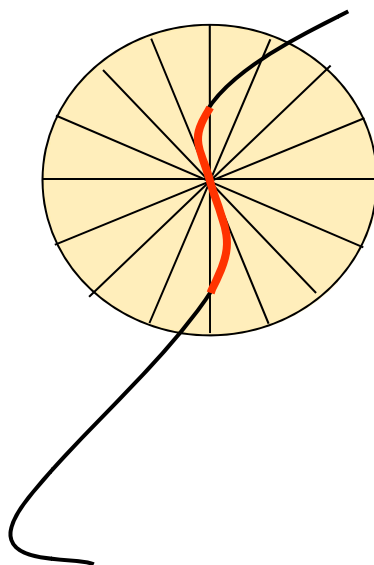
### 3.1.10.2 Practical applications

Three types of S-lines are distinguished:

- S-lines: Pattern recognition of S-lines from a contour line with a certain temperature threshold. The contour line is derived from a smoothed IR image.
- S-lines derived from ‘enhanced’ contour lines: The S-line extraction algorithm is applied on a smoothed IR image where the pixel gradients have been enhanced (cf. section 3.1.1.1.1).
- Large scale S-lines: The same algorithm is used with some modifications in thresholds (on lengths) in order to analyse large-scale wave structures which come with rather weak S-lines.

The identification of S-shaped lines is used for the analysis of the following CMs:

- frontal wave
- upper wave





*Figure 5: Schematic indicating the features being relevant for the S-line detection.*

### 3.1.11 Identification of concave and convex sections of a contour line

#### 3.1.11.1 Algorithm

Let us assume that cloudy ("white") and cloudfree ("black") areas in the satellite image have already been separated by a certain contour line (cf. 3.1.8.1). The remaining requirement for the identification of “white bulges” is the proper definition of a direction of this contour line. The choice here is such that when moving along the isoline in this direction, brighter pixels are always on the left-hand side. Using this convention, one can straightforwardly identify bulges as convex sections of the contour line by considering 3 equidistant points A, B, C on the contour line (where C follows B, which follows A; distance measured along the contour line). The connecting line  $\overline{AB}$  is compared with  $\overline{BC}$ . If the latter line shows an inflection to the left with respect to the former, a convex section is present (analogously, inflection to the right characterises the concave



		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 30/40
---	---	---	--

counterparts). The term “convex” inherently requires a specification of the scale to which the property refers. Hence, an adjustable parameter is the distance of the three considered pixels along the contour line.

Another parameter to be empirically determined is the minimum curvature that has to be exceeded in order to be accepted as clear signal for being convex.

### **3.1.11.2 Practical applications**

Initially the convex pattern of the frontal rear side has been used to identify upper waves. The structure of upper waves ideally shows a convex cloud bulge at the rear side of a cold front induced by a jet streak. Finally the automatic detection of convex cloud pattern were also found to be useful to detect frontal waves since many wave patterns do not exhibit the classical S-shape, either because the cold front part of the wave has no concave rear side necessary for the s-line pattern (see 3.1.10) or because the cold front part of the wave consists of low cloudiness which leads to the determination of the contour line somewhat displaced from the frontal edge.

The “identification of concave and convex sections of a contour line” is applied during the detection of:

- the frontal wave;
- the upper wave.

### **3.1.12 Detection of convective cells**

#### **3.1.12.1 Algorithm**



Convective cells are recognisable by their local brightness maximum (= temperature minimum) and their compact circular or elliptical shape. The steps of their detection are:

- 1) Smoothing of the calibrated IR image in order to filter spurious single bright pixels
- 2) Extraction of the local temperature minimum being colder than a given threshold.
- 3) Around each of these points, a number of concentric circles with prescribed radii are constructed.
- 4) For each circle, it is investigated whether a certain percentage of all pixels lying on the circle exhibit a minimum difference to the brightness of the centre. This criterion helps to eliminate homogeneous and fibrous structures. Each maximum passing the test is marked as the centre of a convective cell.

#### **3.1.12.2 Practical applications**

After the local brightness maxima (temperature minimum) being brighter (colder) than a given threshold are identified, the subsequent steps for the detection of convective cells run with the following parameter settings:

- Concentric circles with the radius  $r$ ,  $0.25*r$ ,  $0.5*r$  and  $0.75*r$  are constructed around each of these points. The choice of the radius  $r$  depends on the size of the phenomena to be identified. Five modules are currently included in ASII. One module is used to detect MCS and the other four to detect CBs and ECs.
- 80% of all pixels on at least one concentric circle have to exhibit a minimum pixel value difference to the brightness maximum in order to trigger a convective cell detection.

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 31/40
---	---	---	--

For the CB detection module, the development stage of the cell is added as attribute on the basis of area-averaged values of the IR development image (see 3.1.2.1): A cell is designated as developing if the decrease of brightness temperature at the coldest part of the cell exceeds 10 K per half an hour; it is decaying if the development value is negative; it is considered as mature otherwise.

### **3.1.13 Dark Stripe detection**

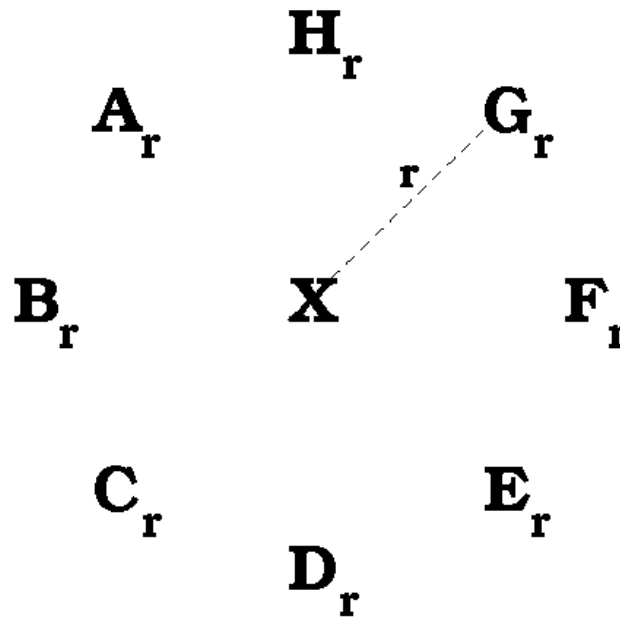
#### **3.1.13.1 Algorithm**

The detection of dark stripes has the following principal steps:

- 1) For each pixel, it is checked whether the pixel is dark enough to be eligible, i.e. whether it is darker than an empirical threshold.
- 2) It is verified for each candidate pixel that it is darker than most of its neighbours. The investigated samples consist of eight surrounding pixels A to H lying on a circle (Figure 6)
- 3) Of these 8 pixel regularly arranged on a circle around the pixel under consideration X, not more than 2 should be darker ("one direction where the stripe comes from, one where it leaves"). The test should be carried out for several radii  $r$  as stripes of varying scale are to be found. The pixel X is considered a potential constituent of a stripe if there is at least one radius where not more than 2 of the surrounding 8 pixel satisfy  $P \leq P(X) + \Delta$  ( $P$  designates the pixel value; note: a brighter pixel has higher numerical value). Setting the required difference in pixel values  $\Delta > 0$ , one can focus on stripes that are distinctly darker than their surrounding.
- 4) It shows that small aggregates of dark pixels can satisfy the conditions checked thus far. However, synoptically relevant dark stripes cover large areas where the checked conditions are fulfilled. Therefore, the algorithm to determine sufficiently large coherent areas (section 3.1.8) is applied with an appropriate empirical threshold on the size, for filtering.

#### **3.1.13.2 Practical applications**

As mentioned above, black stripes in the WV imagery can be a sign for the presence of a nearby jet streak. This information is used as a key parameter for the automatic detection of fibres, of the dry intrusion and of the comma cloud.



**Figure 6:** Configuration investigated during the test for black stripes. The pixel value at point  $X$  is compared with values at 8 selected pixels ( $A_r$  to  $H_r$ ) at several search radii  $r$ .

### 3.1.14 Detection of fibres

#### 3.1.14.1 Algorithm

As fibres can be defined as “white stripes” in the IR imagery, the algorithm for the black stripe detection was adapted to this conceptual model with reversed sign. For optimizing the fibre detection, the following parameters can be set:

- temperature threshold: as jet fibres are an upper tropospheric phenomenon, the temperature of these clouds should be lower than a given threshold;
- critical increase in temperature between the fibre and its surroundings;
- a radius in which the criterion for fibre structure is checked;
- spatial extension of the fibre.

#### 3.1.14.2 Practical applications

The detection of fibres from the satellite image alone cannot exactly indicate the position of a jet streak. The fibre information is combined with the automatic detection of black stripes (see chapter 3.1.13), which are elongated dry zones in the WV imagery. Such dry zones can result from the typical stream field at jet streak level. Dry zones and fibres in juxtaposition give a good indication for the presence of a jet streak.

The automatic detection of fibres is solely used in the detection of jet fibres.



### 3.1.15 Evaluation of texture and spiral structures for CM classification

#### 3.1.15.1 Algorithms

A group of methods is based on the direction of texture elements, i.e. the direction in which the local features of the satellite image appear to be oriented. This parameter can be assessed quantitatively by applying a Sobel filter (see section 3.1.1.1.1) to the image, with subsequent filtering of noise by applying a series of median filters.

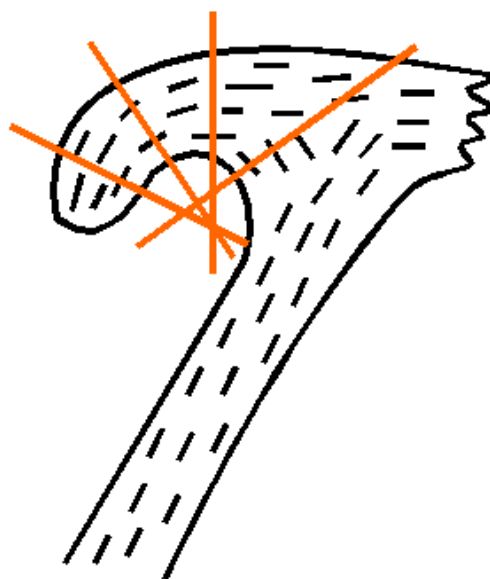
Having the direction at each pixel, it is possible to construct streamlines from this information and, finally, to compute the curvature of these streamlines.

Another evaluation is obtained via the so-called Hough knots. There, straight lines with orientation perpendicular to the respective texture directions, with a certain length and in direction of decreasing greyscale value are drawn. The Hough value is given by the number of rays that pass the considered point. Just as for the curvature, the major signals are obtained in regions of pronouncedly curved structures.



#### 3.1.15.2 Practical applications

The curvature analysis, based on the image texture, is used for the identification of small scale spiral structures like comma features. The typical comma structure, i.e. the small curved cloud band, results in high values of curvature.

The Hough-knot analysis (see Figure 7) is a supplementary method to the determination of curved cloud bands via the curvature. It is used in the detection of occlusion cloud bands. Whilst the curvature analysis flags the cloud structure itself as curved, the Hough-knot method gives the centre of a curved cloud band. This information is exploited with help of a probability function indicating the probability of being near the centre of an occlusion-like cloud band.



**Figure 7:** Schematic of a frontal system with texture elements (black) and Hough lines (orange).

 	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 34/40
---	---	--

### 3.1.15.3 Applications to conceptual models

The curvature analysis is applied to the following CMs:

- comma cloud
- occlusion

## 3.2 PRACTICAL CONSIDERATIONS

### 3.2.1 Validation

The automatic image analysis product has been validated against the manually generated SatRep for 06 and 12 UTC (cf. the PGE10 validation report [RD.1]). The validation period for ASII comprised 8.5 months, starting on 4<sup>th</sup> April 2006 and ending on 12<sup>th</sup> December 2006.

The SatReps were manually generated analyses of the IR and WV image in terms of conceptual models. These analyses have been taken as a reference against which to validate the automatic analysis of ASII. Please consult the full ASII validation report [RD.1] for details, particularly about the definitions of "right" and "wrong" ASII diagnoses.

Amongst the frontal categories in the ASII validation (without model fields), the CF category gives satisfying results with a detection rate of 60% followed by the WF category with 48% and the occlusions with only 35%. In ASIINWP (including model fields), the overall CF and WF detection performs similar with a 60% detection rate for CFs, 48% for WFs. The detection of occlusions works better when model parameters are included (42%).

Concerning the non-frontal categories, lee cloudiness (26%) and jet fibres (12%) are poorly detected. Convective systems show a seasonal variation, being better detected in summer time. Cb clusters (62%) are better detected than MCS (48%), but both are less performing in ASIINWP probably because of inaccurately predicted stability parameters. Convection in cold air (EC and Comma) is less often detected with 22% and 27% respectively. The best detection rate is obtained for the Closed Convective Cells with almost 65%.

Some conceptual models are analysed in ASIINWP only. FI by Jet is detected in 44% of all cases, more than half of the ULL (53%) analysed in the SatRep are detected in ASIINWP. The Upper Wave is detected in almost 40% of all SatRep cases.



It can be summarized that best performances are obtained for frontal categories and small scale convective patterns in the warm sector as well as in the cold air behind a cold front. Amongst the least satisfactory conceptual models are the frontal waves in ASIINWP (detection rate 14%), the jet cloud fibres and the Open Cell Convection (13% in both ASII and ASIINWP).

Since the validation report 2007 [RD.1] was issued, the main changes in PGE10 were

- a) a different scheme of front detection based on temporal changes
- b) a new fibre detection algorithm (prompted by the poor detection found in the extensive validation campaign)

So the current PGE10 version certainly has a better detection rate with respect to fibres. Item a) should have little effect on the detection rates; it rather reduces the false alarms of frontal systems over cold cloudfree land, yet false alarms were not statistically considered during the validation campaign.

PGE10 version 2010 allowed for the first time a flexible selection of the analysis grid. This opened up the opportunity for a new kind of evaluation of the Automatic Satellite Image Interpretation product, namely to assess the uncertainty of the obtained classifications that are due

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD-ASII_v2.0.docx <b>Page:</b> 35/40
---	---	---	--

to numerical (i.e. non-meteorological) factors. Therefore, the following two setups have been examined in the most recent validation report [RD.3]:

- A gradual decrease of the considered analysis region (i.e. removing the outermost grid points without changing the coordinates of the remaining ones).
- ASII analyses on equally sized grids which were, however, slightly shifted against each other.

20 test cases from all seasons and with various synoptic situations went into the statistics.

It was found that generally the detection of conceptual models worked satisfactorily stable when the size of the analysis area has been reduced, in particular for those conceptual models whose detection is primarily based on local pixel analysis (e.g. enhanced cumuli, cumulonimbus and mesoscale convective systems). On the other hand, frontal conceptual models turned out to be more sensitive: if the reduction of the analysis area removes too much of the system, it is not accepted as a front anymore since the PGE10 decision tree explicitly checks for the areal extent as one criterion to be fulfilled. As one of the results, recommendations can now be given on the minimum region size a user should allow so that the analysis for *each* conceptual model over the central region of interest is not negatively impacted by boundary effects. For example, users on the European mainland are advised to use a region extending from Iceland to the eastern part of Europe including Turkey.

The second investigation dealt with small shifts of the analysis grid with respect to a reference grid. Though these shifts were just (approx.) a tenth of the size of a grid point, this has been enough to bring the agreement between the analyses down to 50-75% for small-scale conceptual models which were the worst performers (e.g. Cumulonimbus, fibres...). The most likely explanation for this result is that - when these systems are close to the border of a reference grid cell - they can easily be assigned to adjacent grid points if the grid changes slightly. For frontal categories, such phenomena (that do NOT indicate “unstable” pattern recognition approaches) do not have the same effect (percentage-wise); as may be expected, they are among the best performers with correspondence rates above 80%.

## **3.2.2 Output format description**

### **3.2.2.1 The conceptual model output**

The output format of this product is BUFR. The employed naming convention follows the system for SAFNWC output files as far as possible. It actually adds only the designation of the sub-product, which is: a) based on satellite data only b) based on satellite + NWP data. The following table lists the components used to build the file names.

Field Name	Description	Data Type	Value
Product_id (1)	Identifier for SAFNWC	5 char.	"S_NWC"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id (2)	product name	4 char.	The PGE-specific code "ASII".
Field Separator	Separator	1 char.	"-" (hyphen)
Product_id(3)	Name of sub-product	5 char.	Name of the sub-product coded in the BUFR file. One of the following: "SAT" or "SAT+NWP"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id (4)	Spacecraft identifier	4 char.	"MSGi"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id(5)	Name of processed region	char.	Taken from the region configuration file
Field Separator	Separator	1 char.	"-" (hyphen)
Product_id(6)	Image resolution	5 char.	"VISIR"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id (7)	Date/time of acquisition of SEVIRI image	16 char.	"yyymmddThmmZ"
Field Separator	Separator for extension	1 char.	"."
Extension	Extension of the product name to indicate the type of data	3 char.	bufr

**Table 5: ASII BUFR Output Products Names**

Typical example of resulting filenames for a pair of products is:

S\_NWC\_ASII-SAT\_MSG3\_Test-ASII-VISIR\_20150626T120000Z.bufr,

S\_NWC\_ASII-SAT+NWP\_MSG3\_Test-ASII-VISIR\_20150626T120000Z.bufr.

The BUFR files contain a certain number (in release 2018 of the SAFNWC/MSG package: 20) of "telegrams" which describe the longitudes and latitudes of the grid points where conceptual models are diagnosed. The template for each BUFR telegrams is:

101000 SPECIFICATION OF THE REPLICATION

031002 EXTENDED DELAYED DESCRIPTOR REPLICATION FACTOR

301023 (which expands to:)

005002 LATITUDE (COARSE ACCURACY)

006002 LONGITUDE (COARSE ACCURACY)

000013 ELEMENT NAME, LINE 1

000014 ELEMENT NAME, LINE 2

The following conceptual models appear in the given order in the BUFR output files of PGE10, release 2018. If the strings from ELEMENT NAME, LINE 1 and ELEMENT NAME, LINE 2 are concatenated, these element designations can be reproduced during decoding:

warm front

cold front

cold front under warm air advection

front intensification by jet streak crossing

enhanced cumuli

dry intrusion

cold air cloudiness

mature cumulonimbus

growing cumulonimbus

decaying cumulonimbus

occlusion

comma  
 non-developing wave  
 developing wave  
 upper wave  
 upper level low  
 mesoscale convective system  
 lee cloudiness  
 jet cloud (fibres)  
 embedded convective cell

### 3.2.2.2 The AMV output

An explicit output of the AMV fields (see chapter 3.1.1.2) can be set in the model configuration file (see the Interface Control Document ICD/1 for the External and Internal Interfaces of the SAFNWC/MSG). The output format of these optional products is BUFR. The naming convention is the same as for the conceptual models products, differing only by the designation of the sub-products, which are: a) AMV based on IR 10.8 satellite data b) AMV based on WV 6.2 satellite data). The following table lists the components used to build the file names.

Field Name	Description	Data Type	Value
Product_id (1)	Identifier for SAFNWC	5 char.	"S_NWC"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id (2)	product name	4 char.	The PGE-specific code "ASII".
Field Separator	Separator	1 char.	"-" (hyphen)
Product_id(3)	Name of sub-product	5 char.	Name of the sub-product coded in the BUFR file. One of the following: "IRAMV" or "WVAMV"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id (4)	Spacecraft identifier	4 char.	"MSGi"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id(5)	Name of processed region	char.	Taken from the region configuration file
Field Separator	Separator	1 char.	"-" (hyphen)
Product_id(6)	Image resolution	5 char	"VISIR"
Field Separator	Separator	1 char.	"_" (underscore)
Product_id (7)	Date/time of acquisition of SEVIRI image	16 char.	"yyyymmddThhmmZ"
Field Separator	Separator for extension	1 char.	","
Extension	Extension of the product name to indicate the type of data	3 char.	bufr

**Table 6: AMV Output Products Names**

Typical example of resulting filenames for a pair of products is:

S\_NWC\_ASII-IRAMV\_MSG3\_Test-ASII-VISIR\_20150626T120000Z.bufr,

S\_NWC\_ASII-WVAMV\_MSG3\_Test-ASII-VISIR\_20150626T120000Z.bufr.



Each AMV BUFR file contains a telegram which sequentially holds quadruplets (longitude, latitude, u-component, v-component) of all derived atmospheric motion vectors. The template of each of the BUFR telegram looks as follows:

103000 SPECIFICATION OF THE REPLICATION

031002 EXTENDED DELAYED DESCRIPTOR REPLICATION FACTOR

301023 expands to:



005002 LATITUDE (COARSE ACCURACY)

 NWC SAF	 ZAMG	Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 38/40
--	---	---	---

006002 LONGITUDE (COARSE ACCURACY)

011003 U-COMPONENT

011004 V-COMPONENT

		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 39/40
---	---	---	--

## 4. ASSUMPTIONS AND LIMITATIONS



Pattern recognition methods to analyse the structure of satellite cloud fields are time-consuming processes. This constrained the developers to focus on essential patterns, reducing the number of considered and detected conceptual models. Similar conceptual models, e.g. baroclinic boundary and cold front, are not distinguished and are classified as one category.

It has to be noted, that conceptual models are idealized descriptions for meteorological phenomena. In reality, the variability of these phenomena makes it frequently difficult, even for the trained human analyst, to be sure about the right classification of a certain cloud system. It cannot be expected that the empirical spread defined for a meteorological parameter in such a mechanical system as ASII will be suitable in every instance to yield the right yes/no-decisions. The human analyst in cases of doubt has some mechanisms not accessible to ASII. He/She has the possibility to:

- give little or no weight to features, which do not meet his/her expectation;
- view image loops over a period "as long as necessary". Though ASII takes into account the analysis of the preceding slot, that powerful technique of looping and viewing the genesis of systems is obviously hard to be simulated fully in an objective system.

Low performance must be expected in the proximity of the boundaries of the covered domain. Depending on the pattern recognition modules some require a spacious vicinity for the classification and some need only the close surrounding. Reduced performance occurs also in far-northern regions where the reduced spatial resolution blurs the image features.



		Algorithm Theoretical Basis Document for the Automatic Satellite Image Interpretation Processor of the NWC/GEO	<b>Code:</b> NWC/CDOP2/GEO/ZAMG/SCI/ATBD/ASII <b>Issue:</b> 2.0 <b>Date:</b> 21 January 2019 <b>File:</b> NWC-CDOP2-GEO-ZAMG-SCI-ATBD- ASII_v2.0.docx <b>Page:</b> 40/40
---	---	---	--

## 5. REFERENCES

Bader, M.J., Forbes, G.S., Grant, J.R., Lilley, R.B.E., and Waters, A.J. (Eds.) (1995): Images in weather forecasting. Cambridge University Press, Cambridge.

COST78 (1996): Nowcasting, a survey of current knowledge, techniques, and practice. Phase 1 report. Office for Official Publications of the European Communities, Luxembourg.

Hodges, K.I. (1998): Feature-point detection using distance transforms: Application to tracking tropical convective complexes. *Mon. Wea. Rev.*, **126**, 785-795.

Kidder, S.Q. and T.H. Vonder Haar (1995): Satellite meteorology: An introduction. Academic Press, San Diego, CA, 466pp.

Szantai, A., Désalmand, F., Desbois, M., and Lecomte, P. (2000): Tracking low-level clouds over Central Africa on METEOSAT images. Proceedings, The 2000 EUMETSAT Meteorological Satellite Data Users' Conference, Bologna, Italy, 813-820.

Wirth, A., Jann, A. and Zeiner, B. (2008): On the use of complex empirical orthogonal functions for the temporal interpolation of NWP, radar and satellite data. Proceedings, 2008 EUMETSAT Meteorological Satellite Conference, Darmstadt, 8 - 12 September 2008, EUMETSAT P. 52, ISBN 978-92-9110-082-8, ISSN 1011-3932, 8pp.

ZAMG (1999): Training course: Diagnosis of satellite images with help of conceptual models. CD-ROM (available from ZAMG).

ZAMG (2009): Manual of synoptic satellite meteorology: Conceptual models. Version 6.8. CD-ROM (available from ZAMG).