



IASI

Doc n° IA-TN-0000-2092-CNE

Issue :2

Date :09/07/2001

Rev :

Date :

Sheet :1

of : 21

AVHRR radiances analysis

Inside

IASI FOV's

Prepared by :

F-R Cayla

METEO-FRANCE

Approved by :

G Chalon

CNES



Doc n° IA-TN-0000-2092-CNE

Issue :2

Date :09/07/2001

Rev :

Date :

Sheet :2

of : 21

DISTRIBUTION LIST

X	NOM/PRENOM		SIGLE	BPI	TELEPHONE	FAX
X	BLUMSTEIN	Denis	DSO/OT/SE/IA	2504	05 61 28 26 35	05 61 27 41 72
	CARLIER	Thierry	DSO/OT/SE/IA	2504	05 61 28 20 39	05 61 27 41 72
	CAYLA	François-Régis	SISCLE			
X	CHALON	Gilles	DSO/OT/SE/IA/D	2504	05 61 27 32 27	05 61 27 41 72
	FAURE	Christine	DSO/OT/SE/IA – SECRÉTARIAT	2504	05 61 27 34 60	05 61 27 41 72
	MACIASZEK	Thierry	DSO/OT/SE/IA	2504	05 61 27 41 99	05 61 27 41 72
X	PHULPIN	Thierry	DSO/OT/SE/IA	2504	05 61 27 39 01	05 61 27 41 72
X	PONCE	Ghislaine	DSO/OT/SE/IA	2504	05 61 27 37 10	05 61 27 41 72
X	CHALENCON	Joël	DSO/SG/CT –CONTROLE PROJET	2513	05 61 27 49 58	05 61 28 24 04
X	SEGALEN	Barbara	DSO/SG/CT - DOC PROJET	2504	05 61 27 45 98	05 61 27 41 72
	BRET-DIBAT	Thierry	DTS/AE/INS/IO	2101	05 61 28 18 13	05 61 28 26 92
	BUIL	Christian	DTS/AE/INS/IO	2101	05 61 27 38 97	05 61 28 26 92
	HEBERT	Philippe-Jean	DTS/AE/INS/IO	2101	05 61 27 35 58	05 61 28 26 92
	BOUTEMY	Jean-Claude	DTS/AE/INS/OE	2101	05 61 27 39 71	05 61 28 18 15
	LE NAOUR	Claire	DTS/AE/INS/OE	2101	05 61 28 21 25	05 61 28 18 15
	GUAY	Philippe	DTS/AE/MTE/ME	1416	05 61 28 26 20	05 61 28 29 85
	PRIVAT	Michel	DTS/AE/MTE/ME	1416	05 61 27 37 89	05 61 28 29 85
	SICRE	Jacques	DTS/AE/MTE/ME	1416	05 61 28 21 29	05 61 28 29 85
	PREDINE	Christian	DTS/AE/SEA/AC	1421	05 61 27 36 10	05 61 28 21 62
	DE PLACE	Antoine	DTS/AQ/CA/PC	1412	05 61 28 19 78	05 61 28 13 30
	DESMARRES	Antoine	DTS/AQ/EQE/AM	1414	05 61 28 19 11	05 61 27 47 32
	FRANKE	Frédéric	DTS/AQ/QIS/SC	1415	05 61 28 13 40	05 61 27 42 36
	DANCET	Yves	DTS/AQ/QP/AP	2003	05 61 27 36 85	05 61 27 41 72
	DURIN	Christian	DTS/AQ/SE/QM	1414	05 61 28 14 39	05 61 27 47 32
X	MORENO	Richard	DTS/MPI/PS/TD	1501	05 61 28 25 36	05 61 27 30 84
X	RAYSSIGUIER	Michel	DSO/OT/QTIS/VP	811	05 61 27 36 62	05 61 27 31 67
	DE BOISSESON	Alain	SG/DAF/AC/PE/PP	211	05 61 28 18 75	05 61 28 21 48
	PAPINEAU	Nicole	DP/EOT		01 44 76 77 36	01 44 76 78 68
X	CASSE	Vincent	DPI/EOT	2526	05 61 28 13 85	05 61 27 40 13
	ARNAUD	Michel	DSO/OT/SE/D	2524	05 61 27 30 25	05 61 28 22 26
	GOUDY	Philippe	DSO/OT/D	2524	05 61 28 25 85	05 61 28 22 26
	PIRCHER	Marc	DSO/D	2004	05 61 27 45 22	05 61 27 44 10
	JEGOU	Roger	ESTEC Toulouse	2504	05 61 28 19 30	05 61 27 41 72
X	TOURNIER	Bernard	NOVELTIS		05 62 88 11 15	05 62 88 11 10
	CLAUSS	Alain	ALCATEL BP 99 06322 CANNES LA BOCCA CEDEX		04 92 92 32 14	04 92 92 72 10
X	SCHLÜSSEL	Peter	EUMETSAT Am Kavalleriesand 31 D-64295 DARMSTADT ALLEMAGNE		00 49 6151 807 563	0049 6151 807 656



Doc n° IA-TN-0000-2092-CNE

Issue :2

Date :09/07/2001

Rev :

Date :

Sheet :3

of : 21

DOCUMENT CHANGE RECORD

ISSUE	REV	DATE	Total pages	DESCRIPTION OF MODIFICATIONS
1		30/10/99		Original Issue
2		30/11/2000		Major rewriting including results of algorithm breadboarding

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :4 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------

TABLE OF CONTENTS

- 0. INTRODUCTION 5**
- 1. ASSUMPTIONS 5**
- 2. RESULTS DEFINITION..... 5**
- 3. ALGORITHM 6**
- 4. SAMPLE RESULTS 8**
 - 4.1 INPUT DATA 8**
 - 4.2 CONFIGURATION EXPLORATION 8**
 - 4.3 PROCESSING LOAD..... 10**
 - 4.4 RESULTS ON SELECTED FOVS 12**
 - 4.4.1 NOMINAL PROCESS 12**
 - 4.4.2 USE OF PSEUDO CHANNELS..... 17**
- 5. VALIDATION..... 18**
 - 5.1 COMPARISON WITH CMS OPERATIONAL CLOUD MASK..... 18**
 - 5.2 COMPARISON WITH MAIA 2 RESULTS 21**
 - 5.3 OPEN QUESTIONS 21**
- 6. CONCLUSION 21**

		Doc n° IA-TN-0000-2092-CNE Issue :2 Date :09/07/2001 Rev : Date : Sheet :5 of : 21
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------

0. INTRODUCTION

This document is written in order to give a detailed description of the product radiance analysis inside the IASI pixels. It will describe in section 1 the assumptions on which the AVHRR radiances analysis has been based. It will in section 2 describe precisely the content of the results and in section 3 describe the algorithm used.

It will in section 4 describe the results obtained on two samples of AVHRR overpasses, provided by the Centre de Météorologie Spatiale at Lannion, with the software breadboard of the corresponding OPS functions.

It will, finally, give some preliminary ideas on the validation scheme intended.

1. ASSUMPTIONS

The first basic assumption is that FOV inhomogeneity has to be accounted for accurately when performing retrievals in all but perfectly homogeneous FOV's. It is clear that the partition of the FOV between different boundary surface is necessary for the computation of the outgoing radiances even at infinite resolution. It is also needed to have knowledge of the geometric distribution of these surfaces to precisely take into account interaction with instrument ISRF.

The second basic assumption is that these informations constitute a product at level 1 and are therefore limited to radiances measurements in engineering units avoiding any geophysical interpretation.

2. RESULTS DEFINITION

The description of the results of the radiance analysis is part of the description of the IASI level 1c product as defined in the IASI system specification IA-SP-0000-11-CNE and reads:

“...
b-2 imager data characteristics

The imager data are at this level replaced by a description of AVHRR radiances in the sounder pixels ; these data will be obtained through a detailed multispectral characterisation of AVHRR pixels properties and their separation in a limited number of classes.

FRP-24 : The following information shall be provided (TBC) :

For each line identifications of the AVHRR channels or pseudo channels the radiances of which appear in the sounder pixel description (max 6)

For each sounder pixel :

number of classes identified + 1 (AVHRR pixels not classified) max. 7.

For each group :

**fraction of the IPSF corresponding to the associated pixels,
angular position of its centre of gravity**

mean AVHRR radiances and standard deviation (all channels and pseudo

channels max 6)).

...”

It should be noted that the definition of classes is performed locally (i.e. inside the area covered by one image of the IASI IIS) and that they denote a difference of properties of the AVHRR radiances measured in the area. The discrimination will at maximum separate five classes that will be

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :6 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------

ordered according to the decreasing values of class average radiance in the 11 μ m channel of the AVHRR. It is expected that the number of classes will in general be between 1 and 3, the 5 classes being only reached in multi layered clouds conditions and/or on land sea boundaries.

As class definition is performed over an area larger than IASI pixels, it will frequently occur that a given pixel does not contain all classes, in this case the corresponding fraction will be set to zero but the radiances and standard deviation will give the values determined by the classification over the full area.

This takes into account the fact, that even in simple situations like single cloud layer over sea, the radiance analysis will frequently identify classes associated to partially covered AVHRR pixels. The interpretation of these classes will be facilitated if an estimate of the “clear” radiance is available. As clear AVHRR pixels might not be present in a given IASI pixel, or even in the four while being present in the analysed area it is felt that it is necessary to provide this information to the users.

The angular positions of the centre of gravity of the pixels belonging to one class will be given as azimuth and co-latitude co-ordinates around the interferometric axis, the origin of the azimuth angle being chosen parallel to the scan plane. The average values of associated AVHRR radiances and the corresponding standard deviation will be weighted by the IPSF.

3. ALGORITHM

The central part of the process is a classification algorithm working on the area covered by the IASI IIS at the corresponding scan position (It is roughly. 45x45 AVHRR pixels for scan positions 15 and 16 and its number of lines increase with increasing scan angles).

It is expected to be able to realise a partition of this image using all available AVHRR channels in 5 (TBC) classes leaving a maximum percentage (10% TBC) of unclassified pixels.

As a prototype of the operational software the following method has been implemented:

- a) the mean, covariance matrix and standard deviation of the available channels of AVHRR is computed over the working area
- b) the radiances value are normalised i.e. for each channel corresponding mean value is subtracted and the result is divided by the standard deviation (in order to avoid giving too large a weight to channels dominated by noise the standard deviation is replaced by $N \cdot \text{noise}$ when lower than this value, N adjustable, 10 used right now)
- c) A first guess of class positions is obtained selecting first the pixel whose distance (in the space of normalised radiances noted below DRS) to the mean (0) is the largest, then the pixel whose distance to the first selected is the largest and so on where the quantity maximised is the sum of the distance to the previously selected pixels. This step defines 3 classes positions to which is added the mean position giving a set of 4 initial classes
- d) Then, based on DRS, attribute each pixel to one of the classes and compute mean radiances values for each classes (class prototypes)
- e) An iterative process starts from this first guess of class prototypes. Maximum number of iterations is a parameter of the process. It is currently set to 10.
 - e.1. based on current number of classes, maximum allowable number, current iteration number and maximum number of iterations determine if class separation and or classes aggregation should be performed

		Doc n° IA-TN-0000-2092-CNE Issue :2 Date :09/07/2001 Rev : Date : Sheet :7 of : 21
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------

e.2. if class separation is required :

- e.2.1. select the class to separate :
- number of pixels greater than minimal size
 - spread maximum. The spread is defined as:

$$\frac{\sum_{pixels} \sum_{channels} (R - \bar{R})^2}{N^\alpha}$$

α is a process parameter ,ranging from 1(spread is the variance : number of pixel in the class has no impact on the choice) to 0 (spread is the sum of squared distance to class mean : classes containing many pixels might be selected even if the variance is relatively small)

e.2.2. select inside the class the pixel at maximum DRS from the prototype and the pixel at maximum DRS from the first one

e.2.3. Then based on DRS attribute each pixel of the separated class to one of the 2 prototype and compute mean radiances values for each classes (class prototypes for the 2 resulting classes)

e.3. if class aggregation is required :

e.3.1. Transform the co-ordinates of the centre of classes in the space of radiance covariance matrix eigen vectors.

e.3.2. flag as “not to be aggregated” the classes corresponding to max or min values on the significant eigen vectors (significant if ratio of associated eigen value to maximum one is above a given threshold)¹

e.3.3. compute in this space the inter-classes distance between remaining prototype and select for aggregation the pair corresponding to the minimum, this interclass distance should not be larger that a given threshold.

e.3.4. compute the mean values of the aggregated class

e.4. based on DRS attributes each pixel to one of the classes, if minimum DRS to classes is larger than max allowable DRS the pixel is considered as non classified and compute mean radiances values for each classes (class prototypes); if number of non classified pixels is larger than separation threshold consider them as a class

e.5. Compute the measure of the classification (now defined as the sums of the spreads of the classes) and decide if iteration stops:

- maximum number of iteration reached .OR.
- number of class inferior or equal to max allowable number .AND.
 - ratio of measure to initial value lower than given threshold .OR.
 - .TBD

f). The IPSF originally given in an absolute angular reference around the instrument axis is transformed in weight associated to each AVHRR pixel taking into account the result of the position of the maximum of correlation between the IIS and AVHRR images. These weights are used to compute the means of :

¹ The significance is to be interpreted in term of signal to noise ratio and would imply to compute the covariance matrix on radiance normalised not as described in b) but divided by the noise value for each channel. It should however be kept in mind that this could biases the selection toward contribution of channels having for technical reasons a very high signal to noise ratio as visible channels.

		Doc n° IA-TN-0000-2092-CNE Issue :2 Date :09/07/2001 Rev : Date : Sheet :8 of : 21
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------

Radiances values
 Square of differences between radiances and mean
 column number
 line number

The geometric characteristics, mean column number and mean line number, are converted to angular values.

4. SAMPLE RESULTS

4.1 INPUT DATA

The input data used in this exercise are two fractions of orbit acquired in CMS Lannion during the winter 1999-2000. Each one consists of approximately 4000 lines i.e. close to 15 minutes of acquisition.

They have been provided to CNES in the SATMOS standard format “bandes Master” on a CD which did also contain all necessary software (Fortran sources). It must be noted that the initial delivery did not contain the orbit description file used by the localisation software.

They are:

- a) the night orbit n=24998 on November the 06th 1999 from 03:59:03.749 to 04:13:07.500 The situation is rather cloudy over Atlantic and northern France with clear patches over the Mediterranean sea.
- b) the day orbit n=26063 on .Cloud coverage is nearly ubiquitous even over the portions of north Africa available.

4.2 CONFIGURATION EXPLORATION

Before going to a detailed analysis of the algorithm performance it is necessary to select an algorithm configuration (i.e. set of values attributed to algorithm parameters). The selection has been based on the analysis of the performance of the algorithm in term of its capacity to minimise the distribution spread defined as the sum of the squared distance of each pixel to the centre of the class to which it has been allocated. The initial spread is computed considering that all points belong to a single class the centre of which is the average of all radiances. This reduction is characterised by the ratio of the final to the initial spread which will be called below spread reduction factor and noted SRF (Annex A gives some considerations helping to interpret SRF values).

The first analysis did consider the impact of the number of iteration on the SRF in the case of a maximum number of classes set to 6. The mechanism used to assess convergence and close the iteration has been deactivated and the images were processed imposing 4,5,6,7,8, and 9 iterations.

The results are on the following graph(figure 4.1) showing as a function of iteration numbers the SRF as 50 percentile.

This allow to see that the sensitivity of the SRF to the number of iterations is weak as soon as the number of iteration reaches the value 6 and that there is very little gain to expect going over 9 iterations.

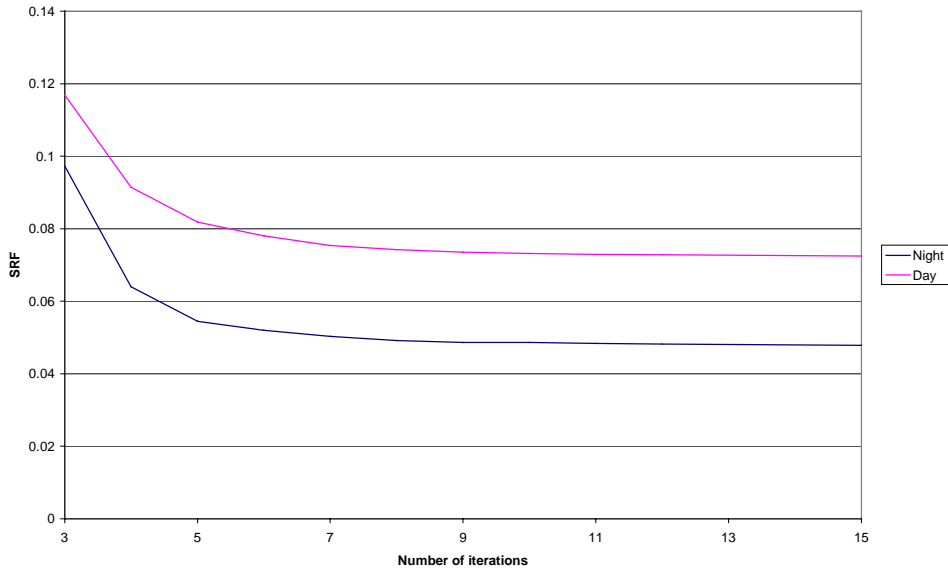


Figure 4.1 Variation of the spread reduction factor versus number of iterations

The second analysis did address the problem of maximum number of classes and reprocessed the two orbits with a fixed number of iterations set to 12 and considered 4,5,6,7,8,9 and 10 classes. The results are on the following graph (figure 4.2) showing as a function number of classes the SRF as mean value or 50 percentile.

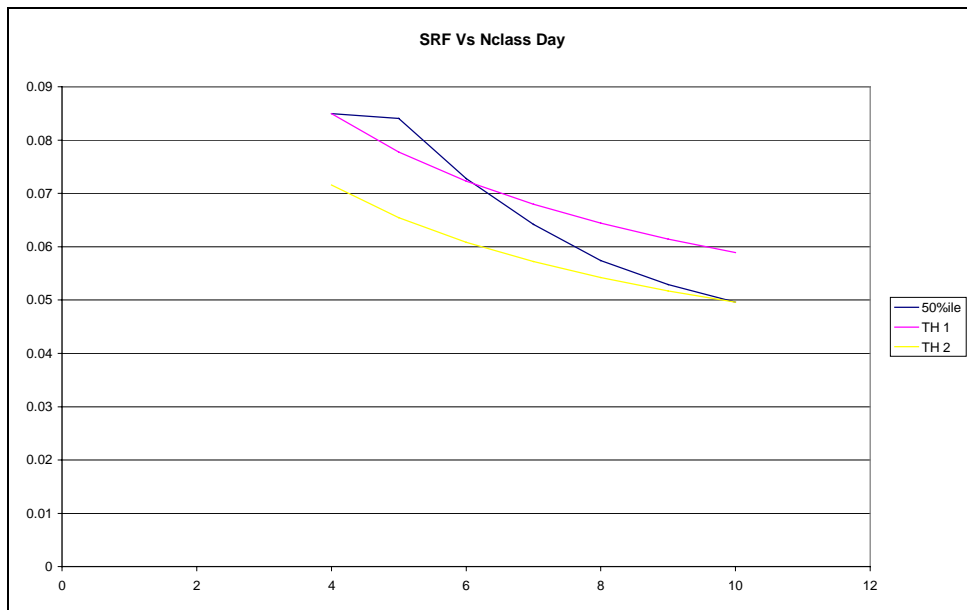


Figure 4.2.a Variation of the spread reduction factor versus number of classes Day

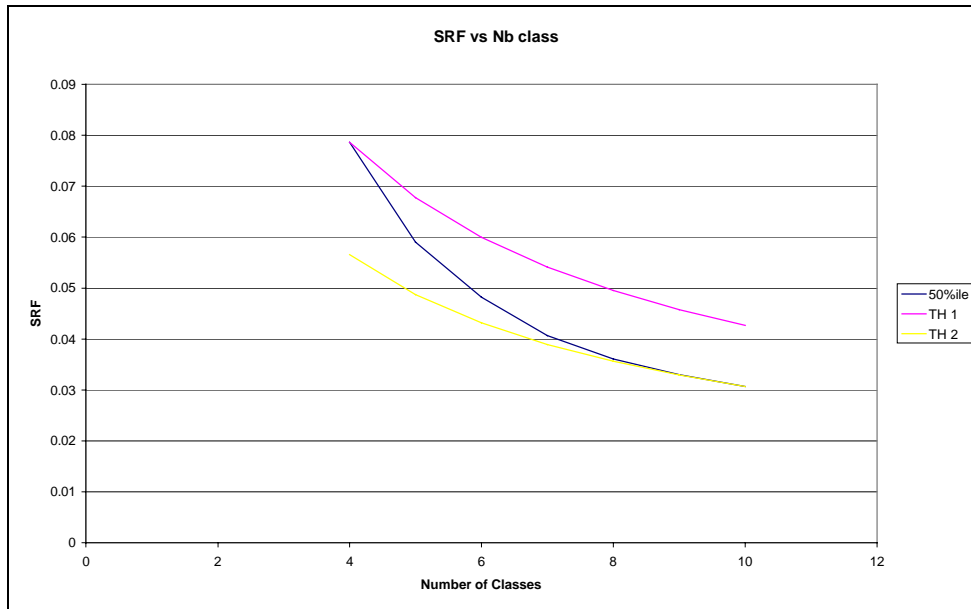


Figure 4.2.b Variation of the spread reduction factor versus number of classes Night

The curve is compared with the theoretical variation corresponding to the purely mechanical effect of partitioning a 5 dimension hyper volume (cf Annex A).In the night case it appears that the SRF decreases rapidly at first and is soon asymptotic to the theoretical variation. This can guide us as to the usefulness of increasing the number of classes describing the radiance distribution and shows that although there is a significant gain when going from 5 to 6 classes any further increase of the number of class brings us in the asymptotic behaviour. The behaviour is not so clear in the day case but the conclusion is basically the same with possibly the usefulness limit shifted to 7 which is not necessarily surprising considering the additional discrimination allowed by the visible channels.

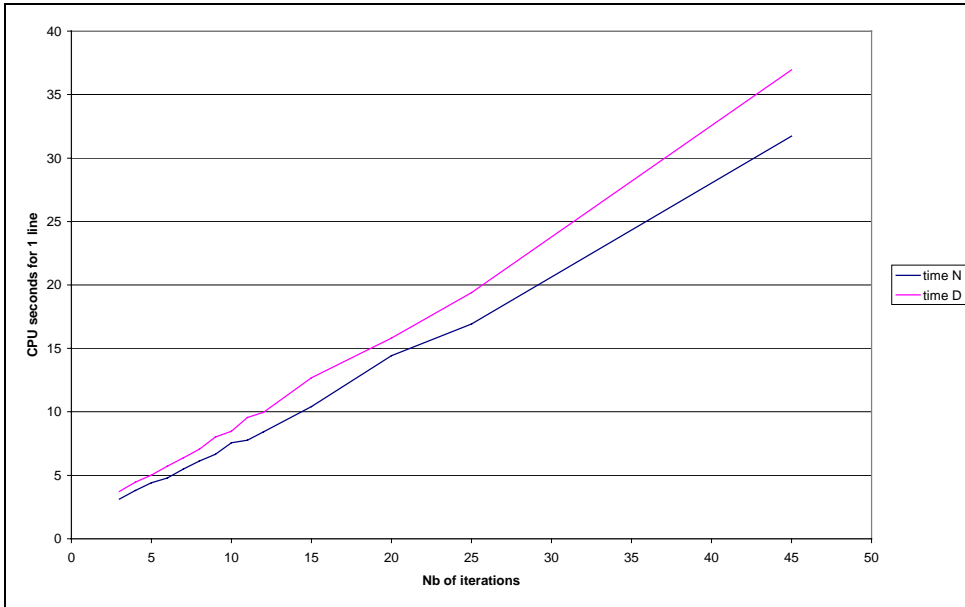
The treatment configuration used as nominal in the generations of the results data file are the following :

Maximum number of classes : 6
 Maximum number of iterations : 9

4.3 PROCESSING LOAD

These treatments have been performed on the standard PC, which is a fairly old one as PCs go (4 year) based on a Pentium Pro at a frequency of 200 MHz. The program used is a standard scientific program in Fortran 77 and the compiler optimisation was used.

The following graphs show the dependency of the CPU time for the processing of AVHRR data associated to one line of IASI data (generated in 8 seconds). The first one shows for a maximum number of class set to 6 the CPU time in second as a function of number of iterations ; the main aspect is the good proportionality to number of iterations. The proportionality coefficient is close to 1 second for 1 iteration.



The following graph shows for a maximum number of iterations set to 9 the CPU time in seconds versus the maximum number of classes



It is there obvious that there is a linear dependency on number of classes but not a direct proportionality ; in this case the slope is close to 0.2 second per class.

The last observation is that the number of channels used has a non-negligible impact that could be evaluated to approximately 10% per channel.

		Doc n° IA-TN-0000-2092-CNE Issue :2 Date :09/07/2001 Rev : Date : Sheet :12 of : 21
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------

Finally I would like to quote the figures corresponding to the selected treatment configuration i.e. a maximum of 9 iterations and 6 classes. These values are 6.66 seconds for 3 channels and 8.02 seconds for 5 channels.

In order to translate these figures in number of operations for one second of incoming data the simple program described in IA-SR-2100-9457-CNE by Richard Moreno was run on the PC and gave a figure of 90 Mflops for the power of this computer.

Taking into account that several improvements of the algorithm have been suggested by the ISSWG, in particular the addition of some pseudo-channels, it appears reasonable to allow a margin for an additional channel bringing the required power to 100 Mflops.

4.4 RESULTS ON SELECTED FOVS

The SRF is a very synthetic parameter which is an indicator of the classification performance but it does not allow to get a feeling as to the relevance of the classification for the purpose of understanding the boundary condition of the direct radiative transfer model.

This paragraph will present the results of this classification on a selection of windows either day or night. It will have two parts showing in the first the results obtained with the classification based on visible radiance and infrared equivalent temperatures (Satmos standard), and in the second the impact of modification of the set of channels.

4.4.1 NOMINAL PROCESS

In order to keep the document manageable, it is necessary to limit the number of case presented to something between 10 and 20. The selection will try to cover:

- a) clear surface properties : homogeneous sea surface, inhomogeneous sea surface, land sea boundaries, mountain area, variable soil properties.
- b) simple broken clouds condition ; high opaque, cirrus, medium, low
- c) multi layer clouds conditions

In the following sub sections a short description will be given on the results obtained for selected scan positions of IASI identified by the line number from the ascending node and the scan position. The associated images are regrouped in annex C to H which also include a listing of the generated product. For each case the images given are first a small extract of channel 5 with incrustation of the IASI sounder 4 pixels then small extracts of one or two AVHRR channels with incrustation of the field of the IIS annotated with line and scan numbers. Another image gives the result of the classification and others show the classification projection on different pairs of channels (bi-dimensional histogram representation). In this annex as in the discussions below AVHRR channels will be designated as Channel n or shortly cn with the usual numbering i.e. :

- 1 Visible
- 2 Visible
- 3 Infrared 3.7 μm
- 4 Infrared 11 μm
- 5 Infrared 12 μm

The following subsections will comment on the individual cases.

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :13 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------

4.4.1.1 CLEAR CASES DAY (CF. ANNEX C)

a) Scan position 50-15 Homogeneous desert

The area is situated in the Sahara over a fairly homogeneous sand area, Dunes patterns are visible in Channel 1 and there seems to change direction between the upper half and the bottom one. The ranges of radiances are low (2 to 3 K in c5). One will note that the classification algorithm separates the two half of the image helped by the different values of c3 for the same values of c5.

b) Scan position 51-14 Inhomogeneous desert

This area is adjacent to the previous one but no longer only on sand. The rocky area shows larger variation of albedo. The classification separate sand and rock and inside the rocky lands the different behaviours results in a geographical partition consistent with the visual analysis.

c) Scan position 85-15 Land sea boundary

This area NW of Sardinia is mostly clear. Sea and land temperatures are very close. Land sea separation is correctly performed based on visible information. Inside the land the classification separate the Clear area from the area round clouds where contamination and also cloud shadows separate another class.

d) Scan position 86-15 Land sea boundary

In this area adjacent to the previous one the algorithm discriminate the island and on it two type of surface. In the same way it separate warmer water bodies.

e) Scan position 86-22 Homogeneous sea

This area of Mediterranean sea is very clear and homogeneous, the radiance range is 1 K in c5. The classification separates the surface in areas consistent with a visual analysis of enhanced images

f) Scan position 92-21 Homogeneous sea

As the previous one this area is fairly isothermal but thermal structure with an amplitude of some tenth of K are present and correctly classified by the algorithm.

g) Scan position 93-21 Land sea boundary

This area covers the Cape of Agde. The classification algorithm correctly separate sea, land and the cold inland water bodies (Etang de Thau, Etang de Vendres). It also separate in a way which seems correct weak thermal patterns over sea (cold coastal waters, warmer current deflected by the cape ...)

h) Scan position 93-22 Inhomogeneous land surface plain

The area covers an inhomogeneous part of the Languedoc containing plain with mostly bare soil (vineyard in winter) and wooded hills. The classification separate the scene in a satisfactory manner. To be noted however the grouping in the same class of a warmer patch in the northeast

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :14 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(bottom left) with the littoral band.

- i) Scan position 93-25 Inhomogeneous land surface mountain, snow
The area covers a part of the Pyrénées. It contains the classical dendritic pattern both in visible and IR. The classification separates the altitude range (temperature) with additional discrimination according to snow cover and state.
- j) Scan position 94-15 Inhomogeneous land surface: mountain, snow.
This area is in the Alps. It looks very similar to the previous one and the classification results are nearly the same.

4.4.1.2 CLEAR CASES NIGHT (CF. ANNEX D)

- a) Scan position 272-6 Land sea boundary
This area is on the coast of Devon. The classification separates clearly sea(red) coastal fringe (green) inland (dark blue); The sequence yellow, light blue, magenta is associated with thin high level clouds.
- b) Scan position 301-25 Land sea boundary
This area is on the Algerian coast. The classification separates land and sea. To note the colder water near the coast, the coastal fringe (dark blue) and the cooler land surface at 20 km from the coast.
- c) Scan position 322-9 Inhomogeneous desert area
Classification is spatially consistent with the images. The curvature of the plot C3 Vs C5 might denote emissivity variations.

4.4.1.3 SIMPLE CLOUD CASE DAY (CF. ANNEX E)

- a) Scan position 73-12 Strato-cumulus
This area is not so simple as initially expected the classification separate a different behaviour of the cloudy layers in top and bottom part with temperature differences reaching 3K. Further separation is based on solar illumination with impact both on visible and C3 channel.
- b) Scan position 73-19 Cumulus over land
The classification separates two classes of Cu with different extension, two type of surface sunny or shaded, and partially cloudy pixels, It is consistent with the visual analysis.
- c) Scan position 101-14 Alto-cumulus.
The classification discriminates two different layers of Ac (dark blue and yellow) with a 6K temperature difference. It also discriminates in the bottom part, another class of cloud characterised by a much warmer equivalent temperature in C3. surface is also well characterised and 2 classes collect partially covered pixels.

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :15 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------

- d) Scan position 102-22 Cumulus, Strato-cumulus
 Correct classification in solid cloud light blue, surface red, and intermediates partly covered. Another type of cloud appears in the bottom left corner (yellow) which due to the linear streaks appearing in adjacent boxes may be interpreted as Ci contamination.
- e) Scan position 124-3 Cirrus
 The classifications separate the Ci according to varying degree of semi-transparency (Clear blue, dark blue, green) over the surface red. In addition it recognises in the upper left corner a different behaviour (yellow, magenta) which can be interpreted as underlying Cu. One can note that the magenta class contains also pixels belonging to the first sequence (clearly apparent in “histo C3 Vs C5”
- f) Scan position 129-5 Cumulus congestus
 Classification according to varying degree of partial coverage.
- g) Scan position 135-29 Stratus
 Classification separate the clouds according to varying degree of partial coverage (Clear blue, dark blue, green). Additional separation at same C5 equivalent temperature (yellow, magenta) denotes impact of different solar reflection (cloud slope orientation with respect to sun direction)
- h) Scan position 137-12 Cirrus
 Classification distinguish the sequence of varying semi transparency (magenta, dark blue) green and blue seems to correspond to different clouds layer on right and left of the Ci band. The yellow is apparently Ci discriminated according to reflectance.
- i) Scan position 138-14 Stratus
 Here also the classification separate a sequence (magenta, dark blue, green) associated to varying partial coverage of the main stratus layer. A different cloud target appears on top (yellow, blue).
- j) Scan position 144-6 Cumulo-nimbus
 The separation is done along a single sequence. As in many of the cases studied above the interpretation of the sequence is facilitated by the contiguity apparent on the classified image it might be interesting to characterise in the product for each class the adjacent ones.

4.4.1.4 SIMPLE CLOUD CASE NIGHT (CF. ANNEX F)

- a) Scan position 235-2 Stratus
 Classification follows a sequence that can be related to variation of partial coverage from red to magenta. However the transition magenta, light blue, yellow denotes apparently a variation of cloud layer properties.

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :16 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------

- b) Scan position 245-4 Strato-cumulus
Classification follows a sequence that can be related to variation of partial coverage from red to yellow.
- c) Scan position 296-26 Strato-cumulus
Classification describe a main sequence (red, green, dark blue) that can be related to variation of partial coverage. In the upper right corner appears another sequence (yellow, light blue, magenta) which is due to varying thickness of overlying cirrus.
- d) Scan position 323-4 Semi-transparent cirrus
This sequence is associated to semi-transparency variation. The high difference in brightness temperature between C3 and C5 for the yellow class seems to denote that the most opaque cirrus are still relatively transparent.

4.4.1.5 COMPLEX CLOUD CASE DAY (CF. ANNEX G)

- a) Scan position 54-29 Cumulus over land
The classification separate the usual sequence (yellow, clear blue, magenta, green and red). It also shows the different behaviour of the shaded land (dark blue) which is darker in visible and colder.
- b) Scan position 64-9 Cumulus developing over Sc ?
The classification is consistent with the spatial structure appearing in the image. The interpretation of the classes in the radiances scatter plot is difficult although one could identify the effects of shadow.
- c) Scan position 68-12 Cumulus developing over Sc ?
Same first comment as above. It is however easier to interpret the sequence dark blue, green, red in term of partial coverage whereas the sequence dark blue, clear blue ,yellow ,magenta seems clearly associated to shading effects and illumination.
- d) Scan position 103-3 Cirrus over ?
It seems that in this case the allocated number of classes is insufficient to describe the fairly complex situation in particular the red class should be separated. This area will be further investigated using pseudo channels.
- e) Scan position 115-12 Cirro-cumulus over Stratus
The classification gives a spatially consistent separation of classes.
- f) Scan position 123-10 Thick cirrus shield overlaying ?
The classification separates two sequences. One (red, green, dark blue, clear blue) seems consistent with varying Ci transparency, the other one (magenta, yellow, clear blue) denotes varying level of solar reflection on the thick Ci.

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :17 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

g) Scan position 126-8 Banded thick Cirrus
 The classification separate two sequences as in the case 123-10. There seems however to be an ambiguous classification for the dark blue class which is spread on the two sequence although it appears clearly separable on the c1 Vs c5 plot.

h) Scan position 143-17 Stratus
 The classification seems dominated by the c3 Vs c5 plot. It denotes small temperature differences (green, dark blue, red) Vs (clear blue, magenta, yellow) and in the sequences varying level of solar contribution.

4.4.1.6 COMPLEX CLOUD CASE NIGHT (CF. ANNEX H)

- a) Scan position 260-13 Cumulus, cumulo-nimbus over strato-cumulus
 The classification shows on one side the strato-cumulus(green) and surface (red)and a sequence(yellow, light blue, magenta, dark blue) associated to cumulus partial coverage. It can be seen on the plot C3 Vs C5 that this last sequence could have been separated in two one corresponding to partial coverage over Sc, the other one to partial coverage over surface.
- b) Scan position 285-26 Cirrus over Stratus
 Two sequences the surface to stratus (red, green, dark blue, magenta) and Cirrus to stratus (yellow, light blue, magenta), the coexisting sequence Cirrus to surface is not separated. One can also note a slightly different behaviour of surface in the upper left corner.
- c) Scan position 296-25 Cirrus over Strato-cumulus
 The classification shows the same two sequences as for scan position 285-26 and as in 260-13 the scatter plot C3 Vs C5 allows to detect that two sequences Ci to Sc and Ci to surface are present but have not been separated.

4.4.2 USE OF PSEUDO CHANNELS

The following configuration of pseudo channels have been explored :

- 0 nominal
- 1 all radiances
- 2 replace the two visible channels by mean and NDVI
- 3 replace two of the IR channels by the difference in equivalent temperature with the third one
- 4 as 3 but the temperature difference is scaled to 280 K
- 5 2 and 3
- 6 2 and 4
- 7 replace the two visible channels by the average of the third IR channel on 3x3 pixels and by the variance over the same 3x3 pixels (Coakley Bretherton approach)

The following comparisons are presented : (TBW issue 3)

		Doc n° IA-TN-0000-2092-CNE Issue :2 Date :09/07/2001 Rev : Date : Sheet :18 of : 21
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------

5. VALIDATION

The validation of this treatment is rather difficult to define. The validation of the steps other than the classification is straightforward.

The classification algorithm could in principle be tested with synthetic data but the project believe that the validation with real data will be a better choice. The difficulty in this case is the definition of the truth. In particular it should be pointed out that, contrary to nominal exploitation of AVHRR data, the absence of the step performing the conversion to geophysical values precludes the use of conventional data as ground truth. In fact this process should be understood as an information compression exercise and tested against a reference algorithm.

The project expected that the validation would compare the classification results on 1 or two AVHRR images with the results given by a reference(TBD) classification algorithm. It is however difficult to find such a reference algorithm. It has been suggested in particular that the classification results should be checked against the operationally generated cloud mask at CMS. It was finally decided to compare the classification results to the operational cloud mask generated at CMS (LUX algorithm) and to the clear and opaque cloud masks generated by the MAIA 2 algorithm developed by CMS in the framework of the AAPP software.

The following subsections will present the results of these comparisons on the selected cases first with the operational cloud mask and in a later issue with the MAIA 2 cloud mask.

5.1 COMPARISON WITH CMS OPERATIONAL CLOUD MASK

The AVHRR data are provided by the Satmos service in CMS with an additional 6th channel containing the results of the operational processing. Inside this processing, pixels radiances are submitted to a family of tests and identified in term of clear or cloudy and, when cloudy, identification of the cloud is provided. This identification selects one of 16 type of targets; they are listed in the following table (T.5.1), together with the colour representation, which will be used in the graph showing the results

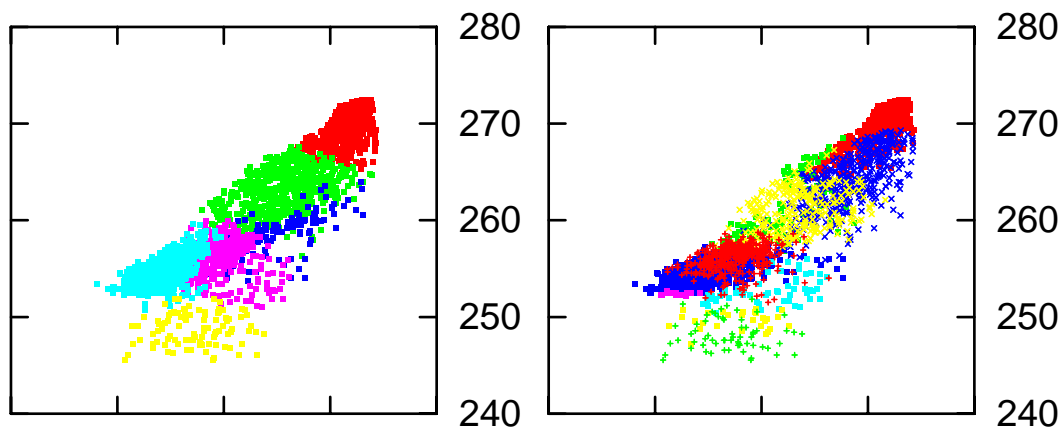
.Table 5.1 Target type code

Code	Target type	Abbreviation tables	Representation
1	Clear	cl	Red sq. dot
2	Very low clouds (stratiform)	vlc	Green sq. dot
3	Low clouds (stratiform)	lc	Dark blue sq. dot
4	Medium clouds (stratiform)	mc	Magenta sq. dot
5	Thin Cirrus	cif	Light blue sq. dot
6	Medium thick Cirrus	cim	Yellow sq. dot
7	Cirrostratus	cs	Red x
8	Thick clouds (Cb, Ns, Ci over Medium clouds)	cb	Green x
9	Broken Clouds, edges	mix	Dark blue x
10	Clear snow or ice	snw	Magenta x
11	Cirrus over low clouds	cil	Light blue x

Figure 5.2 Results of classifications on a scatter plot C3 versus C5

L1c histo c3 Vs c5

Lux histo c3 Vs c5



A last process analyses the behaviour of the histograms of occurrence of the IASI 1c classes in a LUX class. The following table (T.5.2) gives for the same case the number of occurrences.

.Table 5.2 Bi-dimensional histogram between the LUX and the level 1c classifications

Cas	124	3				
LUX\L1c	c1.1	c1.2	c1.3	c1.4	c1.5	c1.6
C1	1140	79				
V1c	1	64	14	29	40	
Lc	0	0	21	210	1013	
Mc	0	0	0	0	15	
Cif	0	0	0	40	11	7
Cim	0	0	0	0	1	28
Mix	120	255	15	13	1	
Cuh	1	153	54	40	7	
Cuhm	0	0	11	247	207	
Cum	0	0	0	0	1	62

The first point to note is the good agreement between the LUX clear and L1c class 1, the difference in column being associated to the LUX mixed class driven by the spatial variance and as such eroding the clear area, the difference in line shows that the clear is also for a small part categorised as class 2 this denoting a threshold effect. At the other end or quite the class 5 is mostly distributed between class Lc and cuhm and there, it is obvious from figure 5.1 that the separation is due more to a variation of the low resolution a-priori information used by LUX than to a difference of properties. The intermediate classes are distributed between the mix and lower clouds cases. For the class 6 it is

		<p>Doc n° IA-TN-0000-2092-CNE</p> <p>Issue :2 Date :09/07/2001</p> <p>Rev : Date :</p> <p>Sheet :21 of : 21</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------

interesting to note that LUX separates it in cumulus and cirrus with no much difference noticeable on the scatter plot C3 Vs C5, I suspect that the classification is associated to C5-C4 values.

In this case there is a good agreement between the two classifications and discrepancies seems minor and can be traced to algorithms peculiarities.

Given the fact that in first approximation the classifications are in both cases driven by a temperature slicing it is not really surprising that they are fairly similar.

Similar graphs and tables for the cases studied earlier are given in Annex I, J, K, L they show in the majority of cases a similarly good agreement and allow to conclude that on this limited sample the L1c classification is separating the areas in a way which can be explained in term of LUX classes. The different number of classes and the absence of a-priori information can explain the inability of L1c to discriminate some targets separated by LUX.

5.2 COMPARISON WITH MAIA 2 RESULTS

The first delivery of MAIA data did not allow a conclusive comparison due to discrepancies in localisation software. Modifications are ongoing.

5.3 OPEN QUESTIONS

The main open question so far is related to the need to provide a geophysical interpretation of the classes in order to be able to specify properly boundary conditions in the direct radiative transfer model. It seems fairly natural to apply the same set of tests to the mean radiances as to the radiances of individual pixels. It is however obvious that it will not then be possible to apply the tests based on local homogeneity characterisation. Specifics efforts will be necessary to check that equivalent results could be obtained trough thresholds applied to the local variance pseudo channel.

The major problem is that this work is outside project area of responsibility and justly so as it is also outside its domain of competence.

It is hoped that knowledgeable members of the ISSWG will endeavour to define a baseline procedure allowing to solve this open question.

6. CONCLUSION

As a first conclusion it can be stated that at this stage the process implemented provides a good descriptions of the scene covered by the IASI pixels. It provides a necessary entry when trying to correct, in the retrieval process, effect of FOV non-homogeneity on ISRF. Its usefulness will however remains limited as long as the retrieval process will not treat explicitly the cloud presence inside the FOVs. It should however be a good way to check that the hypothesis on which are based the implicit cloud clearing schemes are reasonably true within any set of adjacent FOVs.

The second point is that the version presented with this issue of the note is not yet completely frozen and that several improvements suggested by ISSWG are still to be breadboarded and validated.

The final point is that a non trivial work of interpretation of the level 1c radiance analysis in term of geophysical parameters (outside of level 1 scope) has still to be performed. Its first level could well be the use of class average radiances and pseudo-radiances in a classical interpretation algorithm like LUX or MAIA2.