EUMETSAT NWC SAF

SUPPORT TO NOWCASTING AND VERY SHORT RANGE FORECASTING

Optimal use of MTG-IRS spectra on NWC SAF package for Nowcasting purposes

NWC/CDOP3/GEO/AEMET/SCI/RP/IRS_on_CDOP2, Issue 1, Rev.0

31 March 2020

Applicable to close backlog NWC-CDOP2-BL-02

Prepared by Agencia Estatal de Meteorología (AEMET)

REPORT SIGNATURE TABLE

Function	Name	Signature	Date
Prepared by	Miguel A. Martínez (AEMET)		*31 March 2020*
Reviewed by	Pilar Ripodas and Llorenc Lliso, AEMET		*31 March 2020*
Authorised by	Pilar Ripodas, AEMET (NWC SAF Project Manager)		*31 March 2020*

DOCUMENT CHANGE RECORD

Version	Date	Pages	Changes
1.0	*31 March 2020*	77	Report to close backlog NWC-CDOP2-BL-02.

Table of Contents

List of Tables and Figures

1 INTRODUCTION

The Eumetsat "Satellite Application Facilities" (SAF) are dedicated centres of excellence for processing satellite data, and form an integral part of the distributed EUMETSAT Application Ground Segment (http://www.eumetsat.int).

This documentation is provided by the SAF on Support to Nowcasting and Very Short Range Forecasting, hereafter NWC SAF.

The main objective of NWC SAF is to provide, further develop and maintain software packages to be used for Nowcasting applications of operational meteorological satellite data by National Meteorological Services. More information can be found at the NWC SAF webpage, http://nwc-saf.eumetsat.int. This document is applicable to the NWC SAF processing package for geostationary meteorological satellites, NWC/GEO.

1.1 PURPOSE AND SCOPE OF THE DOCUMENT

The purpose of this document is to present a summary of the pioneering activities made at the end of CDOP-2 phase and early CDOP-3 phase related to the adaptation of the previous software for IASI and to explore how to adapt it to future MTG-S IRS data.

These activities was planned in the CDOP-2 working package WP4315E "GEO-sSHAI-MTG-S phase I". As it is shown in this report the activities for this working package was made but due cancellation of second edition of EUMETSAT MTG-IRS Near Real Time (NRT) Demonstration Project and due to lack of time the scientific report has not been written till now.

After the delivering of iSHAI software version 2018 patch 1 for the supporting of GOES-R class and while preparing the NWC SAF 2020 Users Workshop it has been made the effort to write these document to document and to summarize the activities made in CDOP-2 and to close the backlog NWC-CDOP2-BL-02. Activities made during CDOP-3 will be publish in a separated Scientific Report. Thus, on this document is made a summary of the early activities made in CDOP-2 related to:

- Developing of a simple FG regression algorithm with 314 channels of IASI. Simple validation and comparison with the FG regression algorithm of MSG/Himawari/GOES-R/MTG-FCI ones.
- Management of IASI L1 and L2 data and products from EUMETSAT.
- Developing of program readers of EUMETSAT IASI L1 and L2 netCDF files and converter to McIDAS-V compatible netCDF formats.
- Reprojection of IASI L1 and L2 files to NWCSAF region on MSG grid.
- Collaboration on the EUMETSAT MTG-IRS Near Real Time (NRT) Demonstration Project.

All this activities were made as preparation for a future full and synergistic use of MTG-FCI and MTG-IRS data and products from the METEOSAT Third Generation (MTG) on the NWCSAF/GEO software package.

Due to the lack of good synthetic IRS data, because MTG-IRS RTTOV coefficients are not definitive, this document is focused in the use of IASI data and product as proxy of MTG-IRS.

iSHAI (**i**maging **S**atellite **H**umidity **A**nd **I**nstability) is since version 2016 the name of the clear air product of the NWC/GEO software package from GEO imager instrument. GEO-iSHAI product is a heritage of former PGE13 SPhR (SEVIRI Physical Retrieval) product. The name of the product was changed to avoid the word SEVIRI because since release 2018 of the NWC/GEO software package it can be used with other GEO imager instruments different from MSG (AHI on board Himawari satellites and ABI on board GOES-R class satellites).

The name iSHAI was chosen to be coherent with names of NWC SAF products from MTG-FCI and MTG-IRS ("sSHAI) in CDOP-3 and future CDOP-4. Thus prefix "**i**" indicates generated with input data from one GEO **i**mager instrument and "**s**" prefix will indicate generated with input data from MTG-IRS **s**ounder instrument.

In the case of the MTG-FCI software package it will be a continuation of the current MSG NWCSAF GEO package; with almost the same products (cloud mask, iSHAI, etc.) for MTG-FCI Day-1.

In the case of the MTG-IRS hyperspectral instrument it will be developed one optional and independent software package for MTG-IRS Day-2. It will provide three main services/products:

Quick-IRS L1: support of the MTG-IRS L1 activities for local generation of simple like images related products and for generation of input for local generation of NWC SAF or users products. Since EUMETSAT will disseminate by EUMETCast satellite just 300 Principal Components in 160x160 pixels dwell files, this service/product will made:

- First, the Principal Components to BTs spectra reconstruction on every needed dwell file.
- Then for a list of predefined MTG-IRS L1 channels, it will made the combination and reprojection from the needed dwells files to user NWC SAF regions on a geostationary grid. This is needed because the disseminated EUMETSAT dwell files are not reprojected to a fixed grid and just latitude and longitude matrices are included in the dwell files.
- Finally some easy IRS L1 products will be generate. As example they could be generated:
  - MTG-IRS RGB images
  - Normalized IRS-L1 channels in spectral region as CO_2 or WV branches representing the state of the atmosphere at several layers.
  - Simple operations with MTG-IRS L1 as difference or regressions between several IRS L1 images from different wavenumbers.

Thus, users could made the automatic generation of IRS L1 imagery related products or to use them for locally generation of NWC SAF or users products. See References for some examples.

sSHAI_ES: support of the MTG-IRS L2 files disseminated by EUMETSAT Secretariat (ES).

- The combination and reprojection from several dwell files to user NWC SAF regions on a geostationary grid will be made with the disseminated EUMETSAT Secretariat (ES) MTG-IRS L2 products.
- Vertical interpolation from hybrid levels to configurable set of fixed pressure levels.
- Thus, 2D and 3D fields read from the IRS L2 dwells files will be combined and reprojected to users NWC SAF regions.
- Calculation of several nowcasting parameters (TPW, LPW and instability indices) will be made also if needed.

The idea is that ES MTG-IRS L2 fields could be used directly by users in combination with the ones from MTG-FCI fields for same regions. As example comparison of MTG-FCI/MTG-IRS/NWP instability indices.

Also the IRS L2 fields on NWCSAF region could also be used in a future as input to locally generated product. As example IRS SKT field or IRS emissivities could be used as input for MTG-FCI product in future versions.

sSHAI: development of product executed locally by the users. Local generation of NWCSAF MTG-IRS L2 product using light demand CPU algorithms.

- The first one is locally executed light CPU algorithms for retrieval of T, q profiles using as additional input local NWP models for a set of selected dwells. At this moment Xavier

Calbet and Niobe Peinado are developing an algorithm that use kernel ridge regression trained with IASI L1 and NWP analysis of the previous day using a rolling training approach.

- Calculation of nowcasting parameters (TPW, LPW and Instability indices) at dwells.
- Then combination and reprojection from dwells to user NWC SAF defined regions.

There are some remarkable points in the activities described in this report:
a) Use of McIDAS-V tool as main visualization tool and proxy of future user needs and tools.
b) The use of the AEMET tool called PGE00 to made ECMWF 4D interpolation from ECMWF GRIB files to get vertical, spatial and temporally collocated NWP and satellite netCDF files.
d) The PGE00 tool can use the above NWP profiles to get synthetic MSG/GOES-R/MTG-FCI/MTG-IRS/IASI BTs on clear or cloudy pixels. Thus it could be used to get synthetic images or for validation/training purposes. See examples in References.

Some of the results and Figures described in this Scientific Report has been shown in several workshop or communicate to EUMETSAT on private discussions. It will be indicated in each chapter.

1.2 Definitions, acronyms and abbreviations

Please refer to the "Nowcasting SAF Glossary" document in the NWC SAF web for a wider glossary and a complete list of acronyms for the NWC SAF project.

ABI	Advanced Baseline Imager
AEMET	Agencia Estatal de Meteorología Meteorology State Agency (Spain)
AHI	Advanced Himawari Imager
ASCII	American Standard Code for Information and Interchange
ATBD	Algorithm Theoretical Basis Document
BL	Precipitable water in low layer (P_{sfc} – 850 hPa)
BT	Brightness Temperature
CDOP (CDOP-1)	Continuous Development and Operations Phase (1)
CDOP-2	Continuous Development and Operations Phase 2
CDOP-3	Continuous Development and Operations Phase 3
CF	NetCDF Climate and Forecast (CF) Metadata Conventions
CIMSS	Cooperative Institute for Meteorological Satellite Studies (USA)
CMa	Cloud Mask
COTS	Commercial-Off-The-Shelf
CPU	Central Processor Unit
DEM	Digital Elevation Model
ECMWF	European Centre for Medium-range Weather Forecasts
EOF	Empirical Orthogonal Function
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCI	Flexible Combined Imager (MTG)
FG	First Guess
FOV	Field Of View
FOR	Field Of Regard
GEO	Geostationary Satellites
GEO-CMa	GEO Cloud Mask and Cloud Amount
GEO-iSHAI	GEO imaging Satellite Humidity And Instability
GRIB	Gridded Information in Binary Form

HDF5	Hierarchical Data format version 5
HL	Precipitable water in High Layer (500 – 0 hPa)
hPa	Hecto Pascal
HRIT	High Rate Image Transmission
IDL	Interactive Data Language
IR	InfraRed
IREMIS	InfraRed Emissivity
IRS	Infrared Sounder (MTG)
iSHAI	imaging Satellite Humidity And Instability
K	Kelvin
KI	K-Index
km	kilometre
LI	Lifted Index
LPW	Layer Precipitable Water
LST	Land Surface Temperature
MARS	ECMWF Meteorological Archive and Retrieval Facility
McIDAS	Man Computer Interactive Data Access System
ML	Precipitable water in Medium Layer (850 – 500 hPa)
MSG	Meteosat Second Generation
MTG	Meteosat Third Generation
MTG-FCI	Meteosat Third Generation Flexible Combined Imager
MTG-IRS	Meteosat Third Generation Infra Red Sounder
netCDF	Network Common Data Form
NRT	Near Real Time
NWC	Nowcasting
NWC/GEO	Geostationary part of the Nowcasting SAF
NWCLIB	Nowcasting Library
NWCSAF	Nowcasting SAF
NWP	Numerical Weather Prediction
NWP SAF	SAF for Numerical Weather Prediction
LPW	Layer Precipitable Water

PGE	Product Generation Element PGE01 Cloud Mask (GEO-CMa) Product Generator PGE13 SEVIRI Physical Retrieval (SPhR) Product Generator
PW	Precipitable Water
RTM	Radiative Transfer Model
RTTOV	Radiative Transfer for TOVs
SAF	Satellite Application Facility
SEVIRI	Spinning Enhanced Visible InfraRed Imager
SG	Steering Group
SHAI	Satellite Humidity And Instability
SHW	Showalter Index
SKT	Skin Temperature
SST	Sea Surface Temperature
SW	Software
TOZ	Total ozone
TPW	Total Precipitable Water
TM	Task Manager
UM	User Manual
VR	Validation Report
VSA	Visiting Scientist Activities
WV	Water Vapour Channel

1.3 REFERENCES

1.3.1 NWC SAF Applicable Documents

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority. They are referenced in this document in the form [AD.X]

For versioned references, subsequent amendments to, or revisions of, any of these publications do not apply. For unversioned references, the current edition of the document referred applies.

Current documentation can be found at the NWC SAF Helpdesk web: http://nwc-saf.eumetsat.int.

Ref.	Title	Code	Vers
[AD.1]	Proposal for the Second Continuous Development and Operations Phase (CDOP-2)	NWC/CDOP2/MGT/AEMET/PRO	1.0
[AD.2]	NWC SAF CDOP-2 Project Plan	NWC/CDOP2/SAF/AEMET/MGT/PP	1.0
[AD.3]	Request for Changes in the GEO Project Plan for the CDOP-2	NWC/CDOP2/GEO/AEMET/MGT/RP/02	1.0

Table 1: List of Applicable Documents

1.3.2 Reference Documents

The reference documents contain useful information related to the subject of the project. These reference documents complement the applicable ones, and can be looked up to enhance the information included in this document if it is desired. They are referenced in this document in the form [RD.X]

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

Ref.	Title	Code	Vers	Date
[RD.1]	Scientific and Validation Report for the iSHAI Processors of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SCI/VR/iSHAI	1.1	18/12/19
[RD.2]	User Manual for iSHAI Product Processors of the NWC/GEO: Science Part	NWC/CDOP3/GEO/AEMET/SCI/UM/iSHAI	1.1	18/12/19
[RD.3]	Algorithm Theoretical Basis Document for iSHAI Product Processors of the NWC/GEO	NWC/CDOP2/GEO/AEMET/SCI/ATBD/iSHAI	2.2	18/12/19
[RD.4]	Early adaptation of iSHAI v2016 to future MTG-I FCI using 2013 dataset	NWC/CDOP3/GEO/AEMET/SCI/RP01	1.0	31/01/20
[RD.5]	Studies for comparison of NWCSAF/MSG PGE13 SPhR and IASI L2 products	NWC/CDOP2/GEO/AEMET/SCI/RP03	1.0	28/01/16
[RD.6]	Validation Report for "Clear Air Products"	SAF/NWC/CDOP2/INM/SCI/RP/02	1.0	15/10/16

Table 2: List of Referenced Documents

2 INTRODUCTION AND AIM OF THE STUDIES

In CDOP-2 proposal [AD.1] the working package *WP4315E* was foreseen as one activity to do at the end of CDOP-2 (March 2012- March 2017). In Table 3 it is shown the task proposed in this WP.

It must be taken into account that due to the SAFs five year scheduling basis, at the moment when the proposal for CDOP-2 was written in 2011 the schedule of the launch of MTG-S was very different; the launch has been delayed several years and it will be launched finally in CDOP-4.

<table>
<tr><td rowspan="2">WP4315E</td><td colspan="4">Title: GEO-sSHAI-MTG-S phase I (former: GEO-TqPh-MTG-S phase I)</td></tr>
<tr><td colspan="4">Comments:</td></tr>
<tr><td>Start:
Mar-12</td><td>End:
Feb-17</td><td>Effort (mm):
16.00</td><td>Cost:
174.303 K€
(50%Eumetsat)</td><td>Responsible partner:
AEMET</td></tr>
<tr><td>WP Content</td><td colspan="4">Objective:
To elaborate early studies on the use of MTG-IRS spectra on the NWC SAF package for nowcasting purposes.

Tasks:
• Generation of one reduced number of profiles dataset by sampling in the physical retrieval validation profiles dataset for a suitable water vapour and instability training dataset for the METEOSAT full disk.
• Calculation of simulated MTG-IRS BTs spectra dataset for clear air and cloudy (cloud at several layers) conditions from the reduced profile dataset using RTTOV.
• Training and early validation tests of physical retrieval with and without background NWP First-Guess linear regression from full MTG-IRS spectra and with different amount of principal components.
• Calculations of the principal components (PC) for MTG-IRS. Training and early validation tests of physical retrieval with and without background NWP First-Guess linear regression with different amount of principal components.
• Training and early validation tests with the neural network approach.
• Generation of preliminary reports on the use of MTG-IRS spectra on nowcasting by using the NWC SAF package.
• On the other hand, search of optimal RGB combinations through search of optimal simple channels combinations from the analysis of the correlation and error matrices of the simulated MTG-IRS spectra and the nowcasting interest parameter calculated from the profile dataset.</td></tr>
<tr><td>WP Input</td><td colspan="4">RTTOV software.
EUMETSAT software to convert IASI spectra to MTG-IRS spectra
EUMETSAT software for IASI management
McIDAS-V software (HYDRA) for IASI and AIRS management
Neural network and physical retrieval software</td></tr>
<tr><td>WP Output</td><td colspan="4">Reports on the optimal use of MTG-IRS spectra on the NWC SAF package for nowcasting purposes.</td></tr>
<tr><td>WP Interfaces</td><td colspan="4"></td></tr>
<tr><td colspan="2">Interactions with other SAFS and/or Federated Activities</td><td colspan="3">Coordinated effort with EUMETSAT-CAF, with MIST expert group, with NWP SAF and VSA with CIMSS of Wisconsin.</td></tr>
</table>

Table 3: Original formulation of WP4315E in CDOP-2 proposal in 2011.

When the CDOP-2 proposal was written the purpose of this WP was to create a collocated dataset with:

- IASI L1 spectra

- IASI L2 retrieved profiles
- NWCSAF SEVIRI outputs
- ECMWF profiles from several forecast steps and nearest analysis.

This dataset would be later used to make a comparison of the performance from iSHAI and IASI L2 outputs with other truth than radiosounding or NWP. If successful generation of the dataset, a secondary objective was the generation of one training and validation dataset for the different SHAI products; as one alternative to the current use of synthetic NWP dataset.

In this Report the

- In Chapter 3 a test is described that generalize the non linear Fist-Guess regression of iSHAI to the use of BTs from 314 channels of IASI instead of using just the BTs from the 5 channels of MSG. It can be seen the great improvement of the performance.
- In Chapter 4 are described some activities with IASI L1 files.
- In Chapter 5 are described some activities with IASI L2 files. The most important are related to experiences and lesson learned from the participation on the MTG-IRS Near Real Time Demonstration Project.
- Finally some conclusions and some considerations for the future and for the joint operation of NWP, MTG-FCI and MTG-IRS L2 are presented.

3 TEST: FG-REGRESSION WITH 314 CHANNEL OF IASI

The iSHAI algorithm has two main steps: First-Guess (FG) non-linear regression and physical retrieval. The iSHAI algorithm is described in the ATBD [RD.3]. In the iSHAI Validation Reports [RD.1] for MSG/Himawari and GOES-R class satellites and in the Early adaptation of iSHAI v2016 to future MTG-I FCI using 2013 dataset Report [RD-4] for the case of the MTG-FCI instrument, the process to build the First-Guess training datasets, the calculation of the iSHAI FG regression coefficients and the validation is described.

Below the results in a test made to repeat the process using as input to the FG regression 314 IR channels of IASI as a proxy of the future MTG-S IRS instrument performance are shown.

In this test, instead of using the complete iSHAI training and validation 2013 dataset it has been used a reduced and equal weighted iSHAI training/validation dataset extracted from the complete 2013 dataset.

The complete training and validation iSHAI dataset from 2013 year period contains around 3 million pixels; although only those with odd or even position would be taken (as it was made with the imager instrument cases), there are still too many points to perform the IASI simulation with all channels. Then, due to memory and speed constrains in the workstation, it was decided to use only a reduced dataset.

This reduced dataset has been built with 200,000 pixels. Half dataset from sea pixels (100,000 sea pixels) and the other half dataset from land pixels (100,000 land pixels). The reduced dataset has been extracted from the complete 2013 iSHAI training and validation dataset in even positions. In the process of extraction of the pixels, more extraction probability is assigned to pixels less frequently in the total precipitable water histogram, to pixels less frequently in the latitude histogram and to pixels less frequently in the zenith angle histogram.

Then a similar process described in the document [RD.4] was performed using only 314 channels of IASI dataset as a proxy of the future MTG-S IRS instrument performance but with the reduced 200,000 pixels dataset. A first summary was presented at the 2018 Workshop of the Convection Working Group in Ljubljana. The list of 314 IASI channels chosen is the one assimilated by some NWP models at the time of 2016.

Due to the use of a training dataset with reduced number of pixels in this test, it was not possible to perform the spatial analysis of the RMSE that is made in the iSHAI Validation Reports [RD.1] and in the [RD.4] report in the case of the FCI instrument.

In future, the design of the validation programs will be changed so not to need to read and allocate in memory the synthetic brightness temperatures and the profiles and just writing on auxiliary files only the values of the generated parameters. This will allow to generate statistical graphs with the spatial analysis of the error.

The importance of this test is not the early results but to establish a basis of the process and to use the programs for future wider and more significant tests.

3.1 ANALYSIS OF THE PERFORMANCE AT DIFFERENT VERTICAL LEVELS.

In Figure 1, the RMSEs profiles at the 54 RTTOV levels of the q profiles after FG regression step using as input the synthetic RTTOV BTs from 314 IASI channels (as proxy of MTG-IRS) is compared with the RMSEs profiles of the q profiles from ECMWF *t+12* forecast.

In order to facilitate the comparison, in the Figure 2 the equivalent figure for the case of the MTG-FCI instrument generated using the equivalent process but using the complete 2013 iSHAI training and validation dataset is shown. The Figure 2 can be seen in the document [RD.4] where it is described the comparison of the performance of the future MTG/FCI instrument with the performances of previous imager GEO instruments as MSG/SEVIRI, Himawari/AHI and GOES-R/ABI. One difference is that the statistics in document [RD.4] has been made with the odd position on the complete 2013 iSHAI training and validation dataset.

In both cases, the profiles of the ECMWF analysis (*t+00*) from NWP-Hyb datasets (137 hybrid levels interpolated to the 54 RTTOV level) have been considered as the truth. Hereafter, in the figures the meaning of the labels is the following:

1. ***Hybrid t+12***: indicates that the value has been calculated using the comparison of ECMWF *t+12* forecast and the ECMWF analysis *t+00* as truth.
2. **FG**: indicates that the value has been calculated using comparison of the FG result (executing the non-linear regressions of FG step over the satellite synthetic BTs using ECMWF *t+12* forecast as background) and the ECMWF analysis *t+00* as truth.

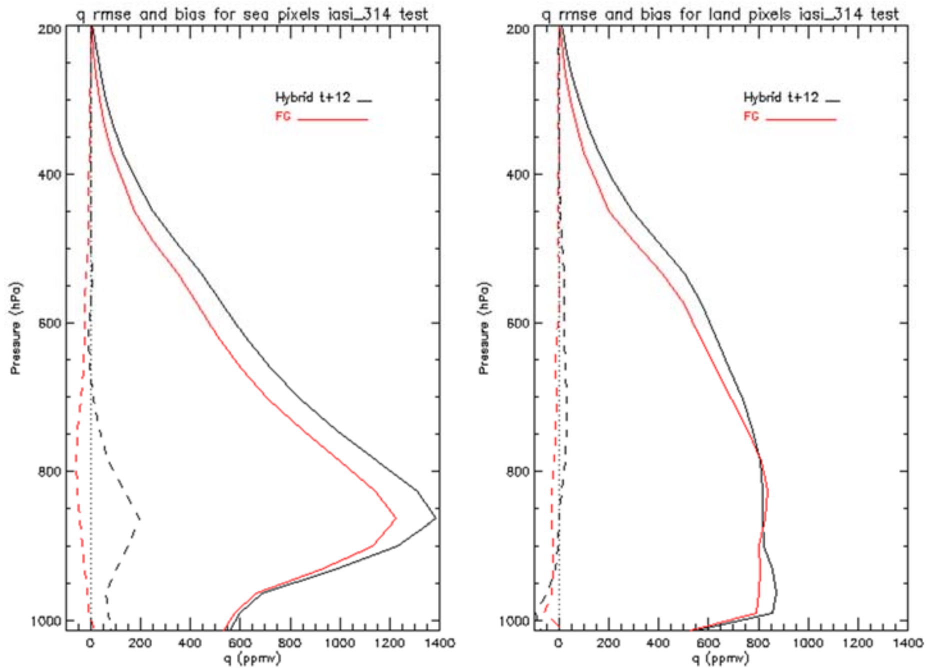


Figure 1: RMSE and bias q profiles (ppmv) from ECMWF t+12 forecast and after FG linear regression step compared with ECMWF analysis (t+00) hybrid profiles ***for synthetic IASI BT_RTTOV 314 channels on the reduced training and validation dataset 2013****. Right) sea pixels, left) over land pixels.*

The statistical values for the specific humidity at mid-levels show better performance for the FG regression than the background NWP model (ECMWF GRIB files on hybrid levels from *t+12* forecast). It can be seen in Figure 1 one strong reduction in the RMSE in the case of 314 IASI channels test.

The retrieved profiles are smoother than the analysis and forecasted ECMWF profiles. For this reason, when RMSE or other statistical parameters are calculated on every pressure level they could be slightly greater than the ones calculated from the *t+12* forecast and analysis ECMWF profiles on some pressure levels; this artefact is more evident on pixels and levels with high spatial and vertical variability as in land pixels on low levels. This has been checked by subjective and interactive comparison of the plots of retrieved profiles versus the analysis and t+12 forecast profiles. This artefact is the cause that in land pixels and in some levels (as near 800 hPa) the RMSE from the retrieved profiles is slightly greater to the *t+12* forecast RMSE.

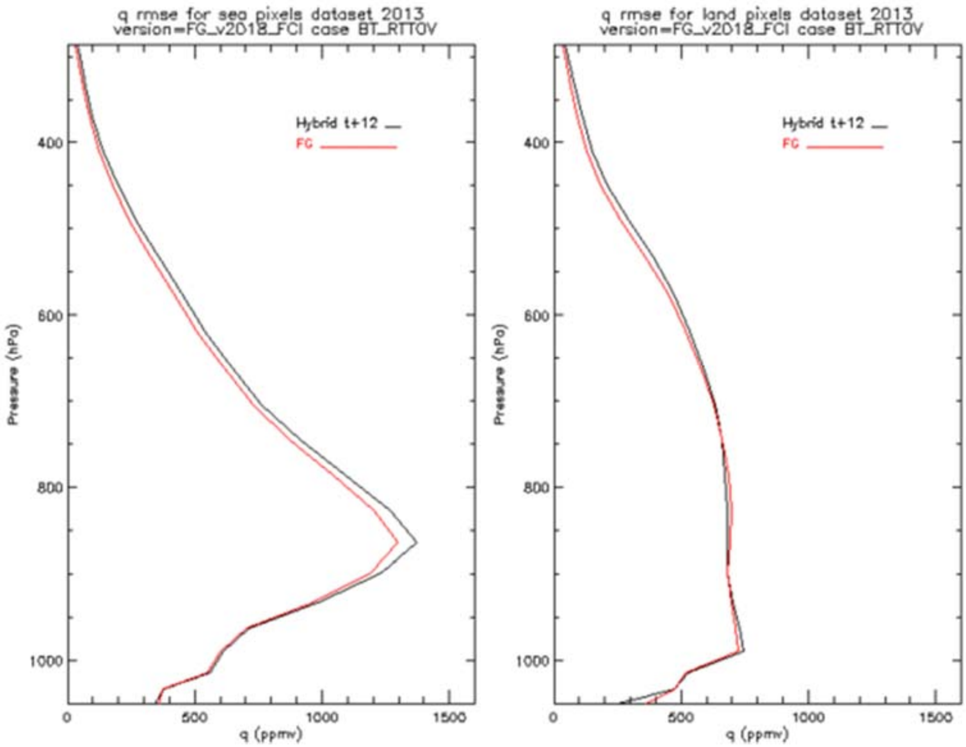


Figure 2: RMSE q profiles (ppmv) from ECMWF t+12 forecast and after FG linear regression step compared with ECMWF analysis (t+00) hybrid profiles ***for synthetic MTG-FCI BT_RTTOV on the complete training and validation dataset 2013****. Right) over sea pixels, left) over land pixels.*

From the comparison of the Figures 1 and 2 it can be seen that the performance in the 314 IASI channel test (Figure 1) is much better than the performance of MTG-FCI test (Figure 2) due to have use a significant greater number of IR channels. This indicates that the future GEO sSHAI will improve significantly the q profile from background NWP. In this experiment just 314 IR channels have been used, since MTG-IRS will have around 2000 channels the use of more MTG-IRS channel or the 300 Principal Components will have much better performance than the MTG-FCI one.

3.2 2D DIMENSIONAL HISTOGRAMS OF LPW PARAMETERS.

To avoid multiplying the number of Figures, only the two dimensional histograms for each one of the Layer Precipitable Water LPW parameters (BL: Bottom layer, ML: Mid Layer and HL: High Layer) and Total Precipitable Water (TPW) parameters calculated from the retrieved profiles from ECMWF *t+12* forecast and after FG linear regression step are presented here.

Note: BL is the precipitable water in a layer between P_{sfc} to 850 hPa. ML is the precipitable water in a layer between 850 hPa to 500 hPa. HL is the precipitable water in a layer between 500 hPa to 0.1 hPa. TPW is the total precipitable water i.e the precipitable water in a layer between P_{sfc} to 0.1 hPa.

For the synthetic BT_RTTOV test, it has been used always as truth the ECMWF analysis denoted here as NWP-Hyb (*t+00*) profiles. In Figure 3 for sea pixels and in Figure 4 for land pixels.

It can be seen in Figures 3 and 4 that statistical values of the GEO-iSHAI parameters on the 314 channels IASI test reproduce the performance suggested by the vertical analysis from Figure 1. The parameters with the largest added value are ML and HL parameters; this fact is due to the contribution of the WV channels.

Other important result is that the 2D histograms of the 314 IASI FG regression parameters does not show significant bias and it would be not needed any postprocessing correction.

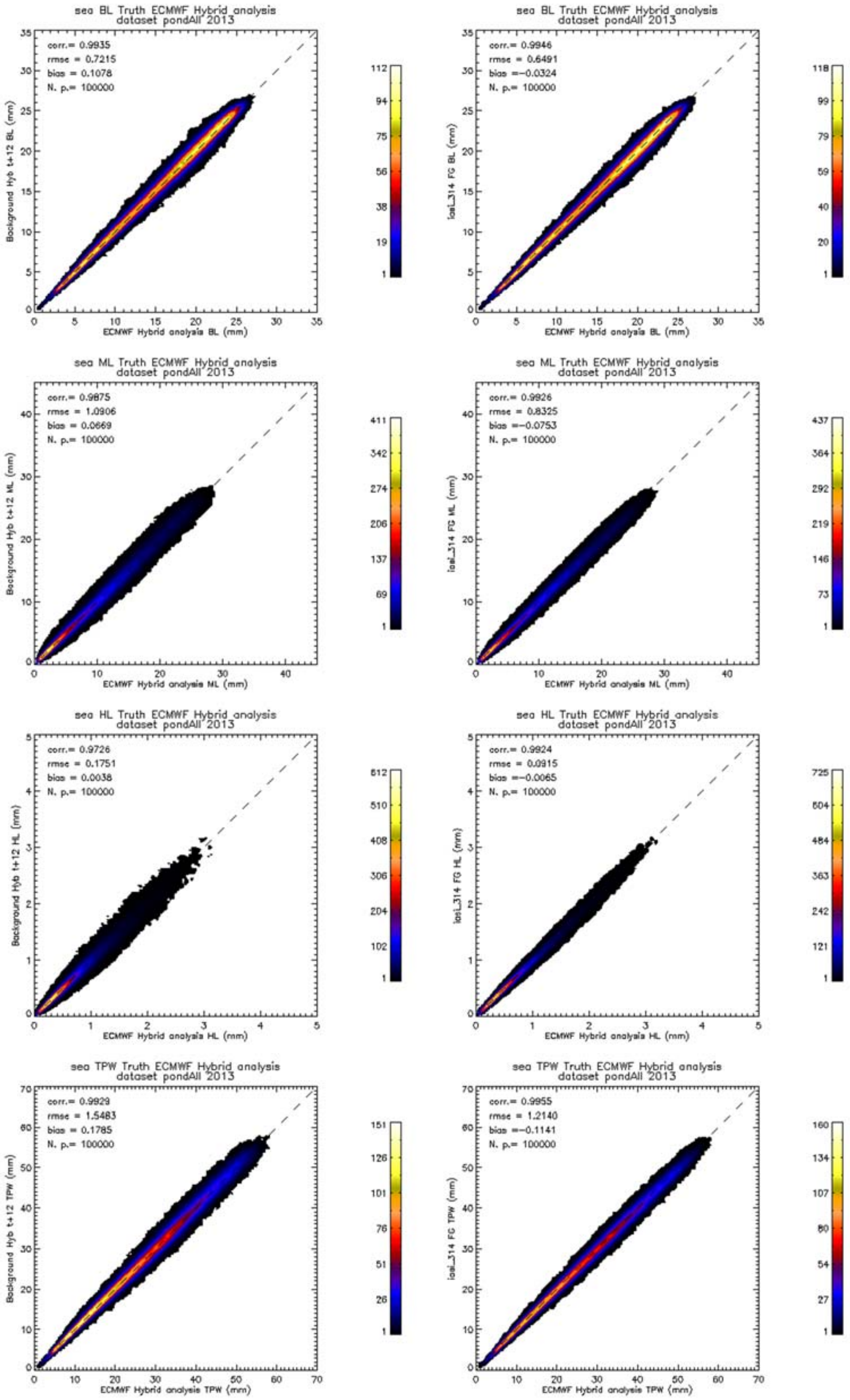


Figure 3: ***BT_RTTOV test*** *for synthetic* ***IASI 314 channels on the <u>reduced</u> training and validation dataset 2013****: LPW and TPW 2D histograms over* ***sea****. From top to bottom BL, ML, HL and TPW parameters. Left) calculated directly from background ECMWF from hybrid profiles from (t+12) forecast, right) calculated after FG step profile using as input BT_RTTOV (t+00). In all cases ground truth calculated from ECMWF analysis (t+00) profiles.*

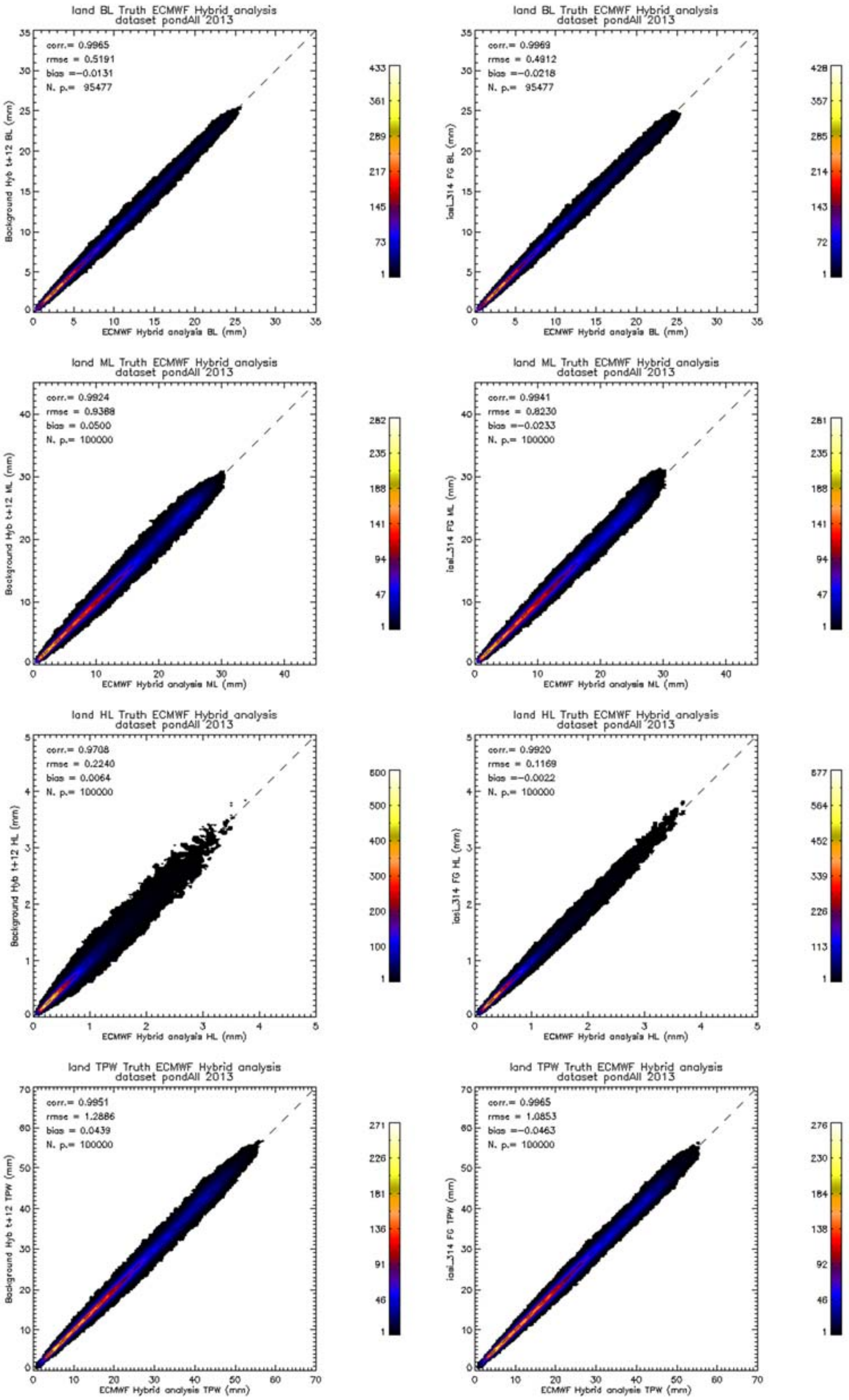


*Figure 4: **BT_RTTOV test** for synthetic **IASI 314 channels on the <u>reduced</u> training and validation dataset 2013**: LPW and TPW 2D histograms over **land**. From top to bottom BL, ML, HL and TPW parameters. Left) calculated directly from background ECMWF from hybrid profiles from (t+12) forecast, right) calculated after FG step profile using as input BT_RTTOV (t+00). In all cases ground truth calculated from ECMWF analysis (t+00) profiles.*

4 ACTIVITIES RELATED TO MTG-IRS LEVEL 1 USE

Some activities related with the use of hyperspectral L1 netCDF files are commented in this chapter. The issues raised and the proposed solutions to contribute to the final objective to get a design of L1 adequate for EUMETSAT and users for nowcasting purposes are also described.

It has been used the sample MTG-IRS L1 netCDF file on Section 4.1 and UMARF IASI L1 on Section 4.2 as proxy of future MTG-IRS L1 files. It has been analysed the netCDF design and contents. Then after check that they cannot be used direct and interactively by the users it was written some IDL/GDL converter programs in order to write hyperspectral netCDF files compatible with McIDAS-V. This approach as proxy of future user's tools has been used to design the service quick-IRS in NWC SAF CDOP-3 proposal.

4.1 PARTICIPATION IN THE MTG-IRS MISSION ADVISORY GROUP AND SCIENCE WORKING GROUP TOPICAL MEETING ON LEVEL 1 PROCESSING

On 22-23 February 2016, EUMETSAT organized a Topical Meeting on MTG-IRS Level 1 Processing with experts, the science working group (SWG) and the MTG-IRS Mission Advisory Group. The main objective of the Meeting was to advance on IRS L1 processing. The author participated as AEMET delegate and as member of the MTG-IRS Mission Advisory Group.

As one of the preparatory activities for this meeting, EUMETSAT distributed a sample netCDF IRS L1 file for a dwell in order to familiarize with the IRS L1 format and to provide feedback on them.

This sample file contains a dwell simulation with synthetic IRS spectra on clear air (without cloud simulations) from NWP profiles on an area covering the Alps and the Gulf of Genoa. The area covered can be seen in Figure 5. It was also distributed in Principal Components (PC) representation on other netCDF file.

Since this 2016 EUMETSAT sample netCDF IRS L1 file could not be used by any standard application, the first activity was to develop an IDL program to read the content of the EUMETSAT distributed sample IRS L1 netCDF file and to convert it to a netCDF file with one internal design that could be used as input to the Hydra component of the McIDAS-V tool.

The converter IDL program read the radiances, longitude, latitude and wavenumber variables and write in a netCDF format compatible with McIDAS-V. Using this converted and compatible with McIDAS-V netCDF files, it is possible to display and use interactively the data in the McIDAS-V tool as proxy nowcasting application of one future user. See as examples Figures 5 to 8 and presentations in References.

This activity demonstrated:

- ✓ The EUMETSAT netCDF design did not take into account the need that EUMETSAT netCDF be compatible with standard meteorological tools of the users.
- ✓ Since users want just to click and use the data then it is important to design and to provide some converters programs that write netCDFs file with internal netCDF structure compatible with users' tool. In this case it has been used as proxy for the process the McIDAS-V tool that could be used just clicking on the converted netCDF file.
- ✓ Since it is difficult to have one convention with standard structure in order the EUMETSAT netCDF file could be open by all the users' tools always be needed some software on the users' side of the EUMETCast to make this conversion.
- ✓ It is not enough to compliance with conventions as CF. It is needed also some standardization in the way the conventions is applied. As example to add dimensions as time, vertical levels, etc.

During the participation on the Topical Meeting on MTG-IRS Level 1 several public and private comments were made. By suggestion of EUMETSAT it was also made a private meeting with the responsible of the MTG file format on EUMETSAT on Level 1 Format Specifications issues. One year later the comments during this activity and others results of the activities on the Chapter 5 was compiled on the "Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files" NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0, Issue 1, Rev.0 on 8th November 2017.

This Technical Note was sent to the responsible of MTG program at this time and upload to the NWCSAF Help desk web as a private documents. Unfortunately the advices made on the Technical Note has not been taking into account and the problems identified persists in the adopted MTG-IRS format. This Technical Note can be seen in Annex II.

The experience after this activity was another reason to propose the Quick-IRS L1 service on the CDOP-3 proposal.

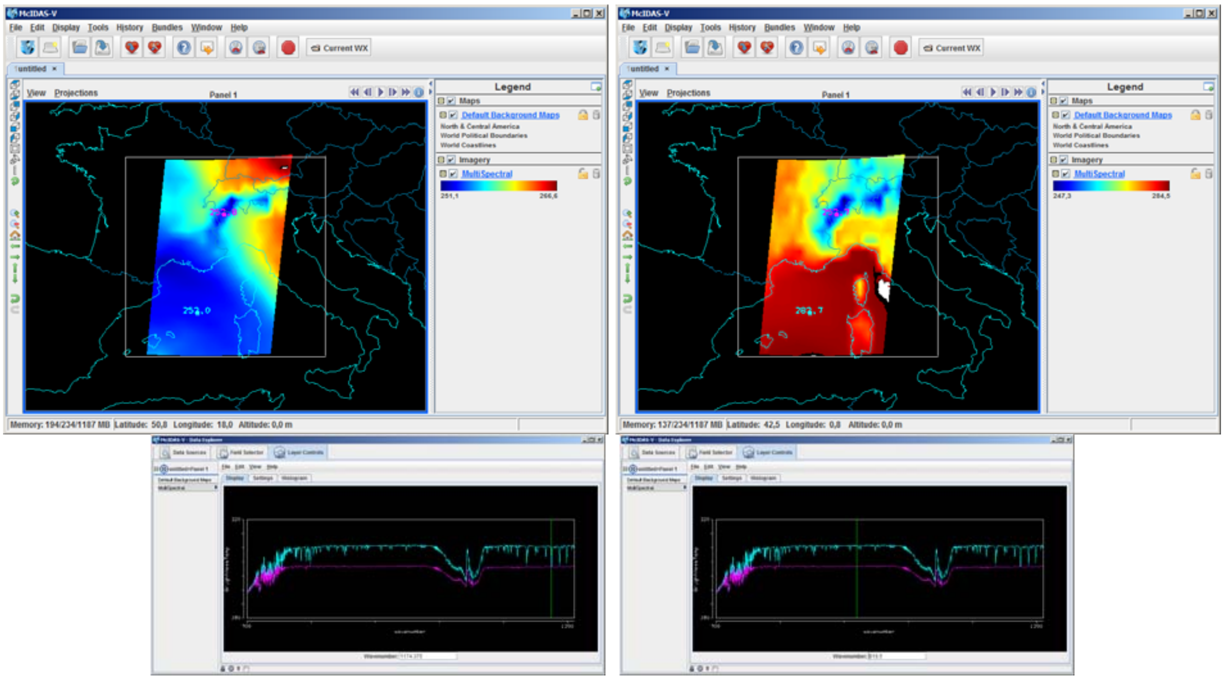


Figure 5: Example of McIDAS-V interactive displays from sample 2016 IRS L1 dwell after conversion of EUMETSAT sample 2016 IRS L1 netCDF files to McIDAS-V compatible netCDF files. (top) brightness temperature at two wavenumbers. (bottom) show the spectra on the pink and blue pixels in the top images on LWIR MTG-IRS band. Users can move interactively the green lines on bottom images to change the wavenumbers of images in the top; moving the pink and blue cross compare the spectra at different pixels.

As one example of the need for this type of software, in January 2020 EUMETSAT distributed other sample netCDF IRS Principal Component (PC) L1 file for a dwell. Unfortunately all deficiencies and problems described in the Technical Note persists. Based on the 2016 IDL procedure it was repeated the process to read the new EUMETSAT MTG-IRS PC L1 file, reconstruct the radiances and BTs spectra from the (PC) and it was converted one more time to a McIDAS-V compatible netCDF file. Examples of the interactive use of this converted file can be seen in the Figure 6.

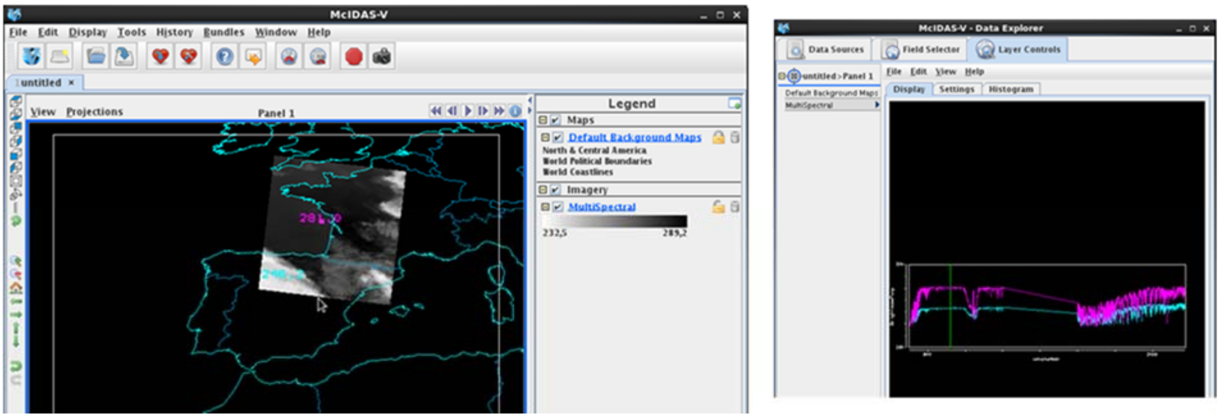


Figure 6: Example of McIDAS-V interactive displays from sample 2020 IRS Principal Component L1 dwell after conversion to McIDAS-V compatible netCDF files. (left) brightness temperature at wavenumber in green line of right image. (right) the spectra on the pink and blue pixels in the left image on LWIR MTG-IRS band. Users can move interactively the green lines on bottom images to change the wavenumber of image; moving the pink and blue cross they can compare interactively the spectra at different pixels.

4.2 ACTIVITIES WITH THE UMARF NETCDF IASI L1 FILES.

For several cases studies it was downloaded from UMARF IASI L1 files in netCDF format to investigate their use in nowcasting. See in References some examples in Martinez [2013, 2015, 2016 and 2018].

One of the key aspects is the need of hyperspectral L1 files to be compatible with adequate exploitation and visualization tools. As proxy of future tools it has been used here the McIDAS-V tool.

On UMARF is possible to download the IASI L1 files in native or netCDF-4 formats. But in both cases the UMARF IASI L1 files cannot be used as input to McIDAS-V.

The problem of UMARF IASI L1 netCDF-4 is that it has been designed just as transcript from binary to netCDF format without taking into account any rule or convention in order to be compatible with any standard visualization tool. This imply that just ad hoc programs or using user developing programs environment as Python can be used to read and use them.

Likely due to Marianne Koenig interest and collaboration with McIDAS-V developers, several years ago the McIDAS-V had special code to read UMARF IASI L1 files on HDF-5 format. But with the time several changes and version updates in the UMARF avoid now to use the UMARF IASI L1 netCDF files in McIDAS-V.

Thus it is needed a converter program. A converter IDL/GDL program was updated to the updated UMARF IASI L1 version using as basis previous prototype as the used in Martinez [2013 and 2015].

At early CDOP-3 phase (end 2017) it was incorporated an important improvement. When the first IASI L1 compatible McIDAS-V netCDF files were used in McIDAS-V the aspect of the images was not smooth and the movement of the mouse was not fully continuous. The reasons were attributed to the noise and low resolution of the IASI images and the huge memory demand to read in memory an IASI array respectively. But after the experience described in Section 5.1.2.2 it was clear the need to reorder IASI L1 pixels depending on the IASI detectors order. As it is described in Section 5.1.2.2 the IASI files are written thinking on the way are observed and to be used in NWP. Thus, they are written not thinking in to be exploited as image arrays. If the use is for NWP or for just plot the data order is not important, but if the use is for display as image field it is crucial to reorder the data to be aligned as continuous arrays. That it is one example of the need of NWC SAF software in the users' side of EUMETCAST to cover the gap between EUMETSAT disseminated files and users' needs.

Some examples of the use with real IASI L1 images can be seen in the figures 7 and 8.

In the Figure 7 it is shown an example of McIDAS-V interactive display of real IASI L1 after conversion to McIDAS-V compatible netCDF files. Users can use move interactively the green lines on spectra images to change the wavenumber of the IASI channel displayed; this allows display IASI channels from different spectral region and due to the difference in the weighting functions travel vertically in the atmosphere.

Also users can move the pink, green and blue probes positions; thus, they can compare interactively the spectra at different pixels of interest (it is also possible to add additional probes).

In the right side of the Figure 7 can be shown the difficult comparison of the full IASI spectra at the three probes (upper right); but with the proposed simplified peak spectra representation (see Martinez [2013 and 2015]) in (bottom right) image the comparison of the spectra can be easily done. In the simplified peak spectra representation, just IASI channel of a set of selected wavenumbers that are relative maxima or peaks in the spectra are written in the McIDAS-V compatible netCDF IASI L1 file. Also it can be made the same with the relative IASI minima wavenumbers (valley spectra representation). Other brain storm (see Martinez [2013 and 2015]) is to add in future users' tool some interactive graphical support to represent at the same time the peak and valley representations. This can of application can be easily configured in the proposed quick-IRS L1 service.

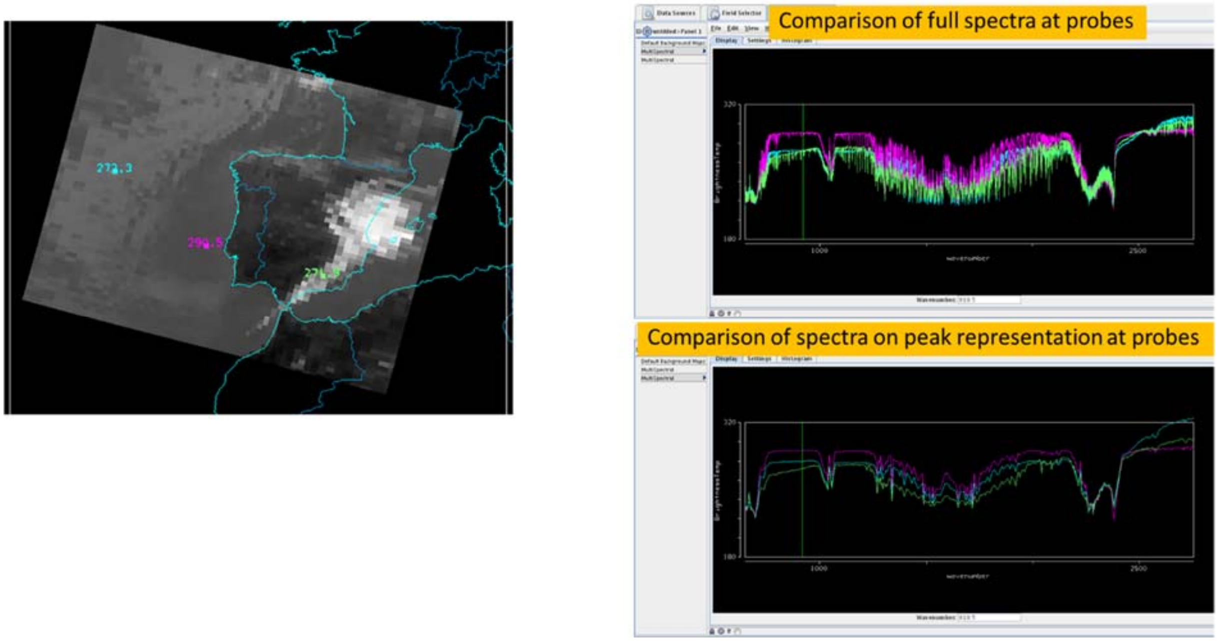


Figure 7: Example of McIDAS-V interactive displays of IASI METOP-B 2016-08-10T10:32:26Z. (left) IASI brightness temperature at wavenumber in green line of right image. (right upper) comparison of full IASI spectra at pink, blue and green probes. (right bottom) comparison of simplified spectra representation with relative peaks representation.

As can be seen in Martinez [2010, 2013, 2015 and 2018] it is possible to generate RGB images with hyperspectral instruments. At least similar RGBs than the ones generated with MSG can be generated. In Martinez 2010 it was made a search of the IASI channels that could be used to generate the RGB equivalent to the MSG RGB ones. In the Figure 8 it is shown the airmass RGB generated with real IASI data for the 10th August 2016 case study. The quick-IRS L1 service will allow to generate a first set of MTG-IRS RGBs images and to investigate the possibilities to generate more in Day-2.

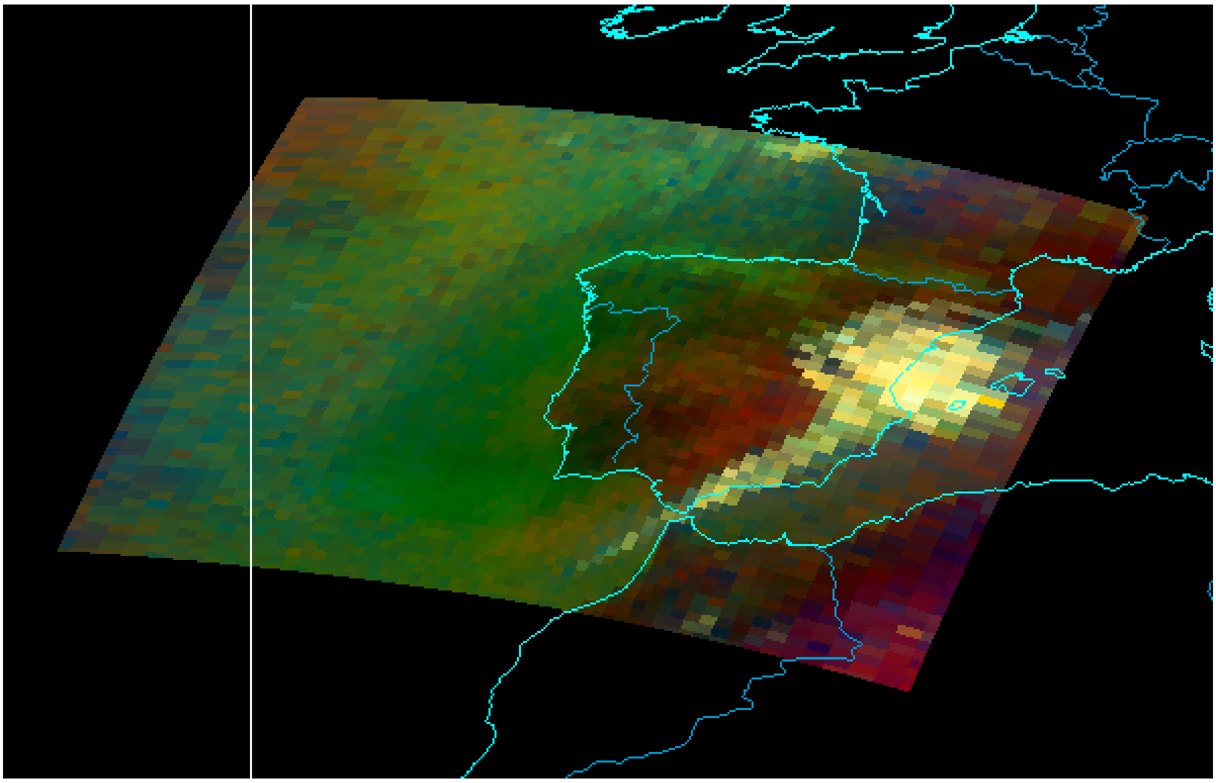


Figure 8: Real IASI airmass RGB METOP-B 2016-08-10T10:32:26Z generated with McIDAS-V interactive display.

4.3 COMMENTS AND PROPOSAL FOR IMPROVING THE MTG-IRS L1 NETCDF DESIGN.

In general the main comments were related to the concern that after the huge EUMETSAT efforts and money on the design, launch and processing of MTG-IRS instrument it is a pity that a poor design of the IRS-L1 format files could block the generalized use by many users if MTG-IRS L1 files if they are not compatible with standard user tools.

Due to this the author suggested to explore and investigate solutions on several issues:

- ✓ Need that the IRS-L1 format follows a standard convention. A practical way to check that the adequate format and convention is followed, is to test the files in a set of state of the art user tools and to have the flexibility to modify the IRS-L1 format till could be compatible with them. It was done but they did not pass. The vision of EUMETSAT is to have a unique format thinking in the EUMETSAT needs.

- ✓ One example of deficiencies detected is to not make correctly de georeferencing. Although the declaration of the latitude and longitude was correctly dimensioned and declared they were not correctly linked with the rest of variables.

  To follow the normal practice in netCDF for users' tools it is needed to establish correctly the georeferencing attributes for any variable. Making a comparison of the declaration and attributes of the variables on the sample IRS L1 netCDF files with the ones in the iSHAI NWCSAF version 2016 files, it was identified a problem in the attributes of the variables. See details in Annex II.

Thus, the issue on how to get a good georeferencing of variables (in the sense to make a correct link of the [latitude, longitude] variables to a variable not in the sense to get a precise navigation) should be revised and tested with other users' tools and future tools developers. In case of need, this attributes could be added or changed later.

- ✓ Add an additional dimension with the wavenumbers. The wavenumber dimension will be used in similar way to the declarations of vertical (or pressure) dimension in the case of 3D arrays of meteorological variables.

  The EUMETSAT UMARF IASI L1 files, add to some variables for each band a set of attributes as the *start_wavenumber, end_wavenumber* and *wavenumber_step*. This fact makes that the user must calculate and reconstruct on memory the wavenumber array. This creates difficulties to make some normal operations as subsetting of some wavenumbers, etc.

  The best solution is to create a coordinate variable containing the value of the wavenumber for each spectral band. Then this wavenumber variable is used as an additional dimension in the declarations of the BTs or radiances arrays. Thus, to declare Radiances(x, y, wavenumbers) is similar to declare T(x, y, pressure_levels) on IRS-L2; in IRS-L2 case it is added one array with pressure levels and is added the pressure as additional dimension.

- ✓ Scale and offset factors. In sample MTG-IRS L1 netCDF files, the radiance_offset and radiance_scale_factor were declared as independent arrays variables.

```
/data/lwir/measured/radiance_offset          H5T_FLOAT [817]
/data/lwir/measured/radiance_scale_factor    H5T_FLOAT [817]
```

  This fact did not follow the standard practice. The standard practice is that this factors be attributes of each variable. The EUMETSAT reason is that this needed due to the radiance range is very different but the price is not to follow standard netCDF codification.

  In the netCDF specification, at this time, scale and offset factors attributes could not be arrays.

- ✓ Need of studies about the use of internal compression of netCDF files always using standard *gzip* compression.

  The variables in the test MTG IRS-L1 netCDF files were not internally compressed. To activate internal compression it is easy in the case of Fortran and C when the netCDF file is created.

  The IDL-8.4 of ecgate machine at ECMWF was used to make the internal compression tests described below. In the case of IDL-8.4 it is possible to activate the internal compression when every variable is defined (this way it can be activated internal compression for one and no compression for other). IDL programs was used to write in one netCDF-4 file just the LWIR radiance, longitude, latitude and wavenumber variables and to test the compression rate using internal compression versus writing uncompressed files. Also, there was made other test executing the external *gzip* compression tool at same compression level.

  As can be seen in Table 4 below the size are similar or better in the case of internal compression with the great advantages that: a) the files can be used after reception avoiding to wait to the end of the external *gunzip* command, b) internal transfer are faster than it would be uncompressed and c) can be directly archived and d) if needed to restore from the archive they are ready to use. Thus, avoid the external compression/decompression could be a great operational benefit for users and EUMETSAT. Test writing LWIR radiance as float (4 bytes) or as short integer (2 bytes) were made.

Size (bytes)	Size of uncompressed files	Compression rate (%)
83877381	IASI_mcIDAS_rad_lwir_float_cxy_20130305000000_20130305000000_N__T_0000.nc	
42046947	IASI_mcIDAS_rad_lwir_int_cxy_20130305000000_20130305000000_N__T_0000.nc	
	Size of files with internal compression	Internal compression rate
43777674	IASI_mcIDAS_rad_lwir_float_cxy_gzip_20130305000000_20130305000000_N__T_0000.nc	52
30928849	IASI_mcIDAS_rad_lwir_int_cxy_gzip_20130305000000_20130305000000_N__T_0000.nc	74
	Size of files with external compression => gzip.exe of file	External compression rate
45701220	IASI_mcIDAS_rad_lwir_float_cxy_20130305000000_20130305000000_N__T_0000.nc.gz	54
34237628	IASI_mcIDAS_rad_lwir_int_cxy_20130305000000_20130305000000_N__T_0000.nc.gz	81

Table 4: Comparison of the size of uncompressed, internal compressed and external compressed files for just with the LWIR radiance, longitude, latitude and wavenumber variables in the case (nchanel, nx, ny).

Related with the internal compression issue is also the issue of the order used to store the 3D arrays. In the Table 4 are shown the results using the order (*nchannel, nx, ny*). In the Table 5 are shown the results using the order *(nx, ny, nchannel)*. It can be seen after this test made with the sample MTG-IRS L1 file that it is better (*nx,ny,nchannel*) but the compression rates could change with the noise of data and further tests should be made.

Size (bytes)	Size of uncompressed files	Compression rate (%)
83877365	IASI_mcIDAS_rad_lwir_float_xyc_20130305000000_20130305000000_N__T_0000.nc	
42046931	IASI_mcIDAS_rad_lwir_int_xyc_20130305000000_20130305000000_N__T_0000.nc	
	Size of files with internal compression	Internal compression rate
29305903	IASI_mcIDAS_rad_lwir_float_xyc_gzip_20130305000000_20130305000000_N__T_0000.nc	35
23122409	IASI_mcIDAS_rad_lwir_int_xyc_gzip_20130305000000_20130305000000_N__T_0000.nc	55
	Size of files with external compression => gzip.exe of file	External compression rate
33485387	IASI_mcIDAS_rad_lwir_float_xyc_20130305000000_20130305000000_N__T_0000.nc.gz	40
26245314	IASI_mcIDAS_rad_lwir_int_xyc_20130305000000_20130305000000_N__T_0000.nc.gz	62

Table 5: Comparison of the size of uncompressed, internal compressed and external compressed files for just with the LWIR radiance, longitude, latitude and wavenumber variables in the case (nx, ny, nchanel).

The use or not of internal compression is a basic decision on the design of netCDF and also it will have a lot of implications not only in the operations but also in science. It must be taking also into account the decisions of archive, operations, etc.

Another recommendation is to avoid use of internal compression using not standard compression functions. As example internal *gzip* compression/uncompression is one standard and usually the majority of users' tools and developing environments (as Python) will have available. But the use of not standard compression as Char-LS could create a handicap to open and read netCDF files on many third party tools for displaying, subsetting, etc.

✓ Several issues and minor bugs also were communicated. It can be seen in Annex II.

These activities and experiences were the reasons to propose the ***Quick-IRS L1*** service and simple IRS L1 products generation on the CDOP-3 proposal.

Quick-IRS L1 will support MTG-IRS L1 activities for local generation of simple MTG-IRS L1 products. Quick-IRS L1 will generate MTG-IRS L1 data remapped to NWC SAF region MTG-IRS L1 for a set of configurable IRS L1 channels. This remapped IRS L1 data will be used later for generation of MTG-IRS L1 imagery (as MTG-IRS RGBs images) or light CPU products (as difference or regressions) or will be used as input for local generation of NWC SAF or other users products over the same regions as the ones defined for MTG/FCI and MTG/LI.

Since EUMETSAT will disseminate by EUMETCast satellite just 300 Principal Components in 160x160 pixels dwell files, this service/product will made:

- Locate the MTG/IRS L1 in Principal Components (PC) format received via EUMETCast needed to cover the user NWCSAF region.
- Made the Principal Components to BTs spectra reconstruction on every needed dwell file.
- Then, it will be extracted a set of predefined MTG-IRS L1 channels and also it will be calculated some configurable light CPU products (could be regressions, difference of channels, normalization of BT of several channels, etc.)
- Then, it will made the reprojection and combination of the fields above from the needed dwells files to user defined NWC SAF regions on to MTG-FCI projection (geostationary grid). This is needed because the disseminated EUMETSAT dwell files are not reprojected to a fixed grid and just latitude and longitude matrices are included in the dwell files. Also there is a slight spatial displacement between LWIR and MWIR bands.
- Finally some easy IRS L1 products on NWC SAF regions in MTG-FCI projection will be written in netCDF files.
- Then some simple MTG-IRS L1 products will be generated. As example:
    - MTG-IRS RGB images
    - Normalized IRS-L1 channels in spectral region as CO_2 or WV branches representing the state of the atmosphere at several layers; allowing to display normalized BT on channels peaking deeper and deeper in the atmosphere
    - Images files with simple operations with MTG-IRS L1 as difference or regressions between several IRS L1 images from different wavenumbers.
    - Use of set on IRS L1 with interactive tools as McIDAS-V to interactively display spectra or subset of spectra over a pixel.
    - Display in combination of other MTG-FCI or MTG-IRS L2 products suitable for use in nowcasting, e.g. stability indices.

Thus, users could make the automatic generation of IRS L1 imagery related products or to use them for locally generation of NWC SAF or users products. See some examples on References.

5 ACTIVITIES RELATED TO IMPROVE THE USE OF EUMETSAT MTG-IRS LEVEL 2 IN NOWCASTING

In this Chapter are described some activities with IASI L2 files. The most important are related to experiences and lesson learned from the participation on the MTG-IRS Near Real Time Demonstration Project.

5.1 PARTICIPATION ON THE EUMETSAT MTG-IRS NEAR REAL TIME (NRT) DEMONSTRATION PROJECT

The participation on the EUMETSAT MTG-IRS Near Real Time (NRT) Demonstration Project was initiated at end 2015 after an invitation of at that time responsible of MTG-IRS, Stephen Tjemkes, and the responsible of the MTG program Rolf Stuhlmann.

It was decided the NWCSAF AEMET participation due to:

- During the MTG-IRS NRT demo project, IASI data from METOP-A and B over Europe were processed in NRT mode with the at that time prototype for EUMETSAT MTG-IRS L2 algorithm; based in physical retrieval using as Fist-Guess (FG) the ECWMF model.
- Thus, MTG-IRS NRT demo project products could be considered as a source of proxy MTG-IRS data to build a training dataset in accordance with proposed AEMET NWC SAF CDOP-3 activities for MTG-IRS preparation.

The proposal for AEMET participation in the EUMETSAT MTG-IRS NRT Demonstration Project can be seen in Annex I. After a coordination Webex meeting in 11th December 2015 the MTG-IRS NRT Demo Project started. During the MTG-IRS NRT Demonstration Project there were some monthly coordination Webex meetings. It was also used Basecamp as collaboration tool in the project.

The NWCSAF AEMET contribution in the project was based in the download, read of MTG-IRS proxy data files inside of the AEMET region of interest (see Figure 1 of Annex I), the combination of several files in the region and the writing in a netCDF format compatible with McIDAS-V tools (used as proxy of future meteorological tools). This way the MTG-IRS proxy data could be used and compared with other L1 and L2 products (NWP, NWCSAF iSHAI products from MSG, sSHAI from IASI prototype and EUMETSAT-ES UMARF IASI L2 v6.2 products). The process at AEMET was similar to the proposed ***sSHAI_ES*** service for the future disseminated EUMETSAT Secretariat MTG-IRS L2 products and the integration on the NWCSAF GEO software package.

5.1.1 DESCRIPTION OF AEMET PROCESS OF THE MTG-IRS PROXY DATA

The original idea was to distribute the MTG-IRS proxy data files using EUMETCAST Terrestrial by GEANNT Internet Fast network. But several of the participants we hadn't EUMETCAST Terrestrial access. By this reason the products were made available in NRT stream using a restricted access EUMETSAT ftp server.

In the ftp server there was a set of directories by processing region. In each directory there were files with 5 minutes of IASI orbit as proxy of the MTG-IRS segmentation on dwell files.

At AEMET side it was developed one Python program triggered by *crontab* that every half hour check at the EUMETSAT ftp server in the directory with IASI files over Europe region if there is any new files and download them.

Note: CrIS data on board of NPP-Suomi satellite was also processed in the reprocessing mode. But it was not used because the CrIS datasets has 9 detectors that rotate 30º along the track and it is difficult to build image fields.

Once download the IASI netCDF-4 files in the preprocess chain, it was made:

- The IASI MTG-IRS proxy netCDF-4 files were open and read data using GDL HDF-5 commands (see explanation in Section 5.1.2.1).
- The array of data were saved on binary files using the GDL *save* command.
- Data were *restored* on IDL and the IASI lines with 120 pixels were reorder to correct IASI detectors order (See Section 5.1.2.2).
- The data were reform as an array of structures set of variables for **FG**, **IASI** and **other**.
  - In **FG** structure are stored all variables with name including prefix '*PRIOR_*' together some variables as longitude, latitude, zenith angles, julian date, etc
  - In **IASI** structure are stored all variables with name including prefix '*POSTERIOR_*' together some variables as longitude, latitude, zenith angles, julian date, etc
  - In other structure are stored the rest of variables. …
- The temperature/q/ozone profiles from hybrid levels were interpolated to a set of configurable pressure levels (See Section 5.1.2.3).
- Some nowcasting interest parameters as TPW, LPWs and instability indices were calculated
- Writing of the arrays on to McIDAS-V netCDF-3 files. An extra dimension for time was added. This time dimension was dimensioned with two times positions. The arrays from the FG are written in first position with time 0 second and the physical retrieval are written in the second position with time 1 sec. One example of utility the time dimension for easy comparison of the changes between FG and physical retrieval can be seen in Figure 14.

The first part of GDL/IDL chain acts as converter program from netCDF-4 to netCDF-3 compatible with McIDAS-V on IASI coordinates. Thus, the pre-process of MTG-IRS proxy netCDF-4 files was made with the limited GDL and the core of process remains made with IDL.

These McIDAS-V netCDF-3 files for every 5 minutes IASI can be used for interactive and batch display by users. But in second part of chain they were combined several in one set for the orbit segment covering the region of interest.

Then it was started the developing of a reprojection program from IASI to MSG regions. This allowed a wider and easy comparison of the ECMWF, NWC SAF and IRS-demo data as proxy of future use by users and for the generation of validation datasets. Thus the third part of chain the IASI orbits could be remaped to the NWCSAF region of interest.

5.1.2 Some interesting details of the processing chain.

5.1.2.1 Use of IDL/GDL HDF-5 functions to read netCDF-4

Due to that at AEMET the IDL version is very old (IDL version 6.2), the netCDF-4 files can not been read with it. Also, the solution to use some external tools (like CDO) to convert the MTG-IRS NRT Demonstration Project netCDF-4 files to netCDF-3 did not worked due to some deficiencies in the MTG-IRS NRT demo netCDF-4 design and that some values were not well initialised.

Since at this time at ECMWF in *ecgate.ecmwf.int* machine IDL was available with the modern version 8.4, it was made the process to upload to ECMWF the early MTG-IRS NRT demo netCDF-4 files and the early tests with them were made on *ecgate.ecmwf.int* machine.

But just when the first IDL 8.4 prototype was ready it was received an email from the ECMWF (Xavier Abellan) informing that in the next update of the *ecgate.ecmwf.int* machine since February 2016 it will be removed the access to IDL by non ECWMF staff and suggesting to try a free version of IDL called GDL.

GDL is very limited and it is far to be as powerful as IDL but the limitations of GDL were important to discover some very interesting facts.

In the initial test of GDL it was discovered that the GDL netCDF functions has a bug and that GDL cannot read data from any netCDF-4 files. Looking for a solution it was tested the use of the GDL HDF-5 functions to read the data from the MTG-IRS proxy netCDF-4 files and it worked well. This is possible because **a netCDF-4 file is indeed a reduced variant of the HDF-5 family**.

Below in Table 6 can be compared a sample codes with IDL8.4 functions to open and read data with the netCDF or with HDF-5 way. It can be seen that read of data with HDF-5 function is more direct (if the name of the field is known) and easier than with the netCDF functions.

Using netCDF commands	Using HDF-5 comands
.. name="IASI_BSC_1C_M02_20151218140853Z_20151218141157Z_N_O_20151218145151Z.nc" cdfid = NCDF_OPEN(Filename) skt_variable_name = strlowcase('POSTERIOR_SURFACE_SKIN_TEMPERATURE') varid = NCDF_VARID(cdfid, skt_variable_name) NCDF_VARGET, cdfid, varid, skt ..	.. name="IASI_BSC_1C_M02_20151218140853Z_20151218141157Z_N_O_20151218145151Z.nc" h5id = H5F_OPEN(name) skt_variable_name = strlowcase('POSTERIOR_SURFACE_SKIN_TEMPERATURE') Dataset_id = H5D_open(h5id, skt_variable_name) skt = h5d_read(Dataset_id) ..

Table 6: Comparison of IDL codes to read the posterior SKT field (i.e. SKT retrieved with IASI).

Since GDL is free (no license need) it was installed at AEMET and then GDL was used with the HDF-5 functions to read the data from the MTG-IRS proxy netCDF-4 files and write the data on binary files that can be easily *restore* with IDL or GDL.

Note: IDL and GDL have the command save that write on a binary file a list of variable in memory in a file with a description of the type and dimensions. The command restore allows to read in memory from the binary file the saved variable without need of allocation. Other languages as numpy and R has equivalent functionality.

5.1.2.2 Reorder of data to correct the order of the IASI detectors.

One key point was the reorder of data to correct the IASI detectors order.

When used the early McIDAS-V netCDF-3 was clear that there was something wrong. Due to a comment of Xavier Calbet it was noted that all EUMETSAT IASI files are written using a magic number of 120 elements with a row of IASI observation. Since there are 4 IASI detectors the natural choice is reform the data to 2 lines with 60 columns. But instead the IASI detectors to be numbered as [[0,1],[2,3]] they are numbered as [[2,1],[3,0]] and this creates a problem as can be seen below.

By this reason is needed before to reorder the array with IASI 120 pixels of a line to arrays that could be displayed as images taking care to the IASI detector order. In the table below is represented one line of 120 profiles in a MTG-IRS proxy array and are colored the four detectors (0^{th} yellow, 1^{st} green, 2^{nd} blue, 4^{th} pink)

0 1 2 3 4 5 6 7 ……………………………………………………………119

After reordering the line of 120 profiles it must be converted to 2 lines of 60 profiles and the profiles must be ordered as in table below. See the color sequence to get the IASI detector sequence.

2	1	6	5
3	0	7	4

After making the above reorder of the IASI detector McIDAS-V netCDF-3 could be used without any problem at it was also possible to create interactively sounding diagrams.

As can be seen the EUMETSAT IASI L1 and L2 files are written thinking on the way are observed but not a natural way to be used for users.

5.1.2.3 Interpolation from hybrid levels to a configurable set of fixed pressure levels.

Since other objective of the MTG-IRS Demo project was to investigate also the use L2 products in NWP assimilation on limited area model the temperature, humidity and ozone profiles on the MTG-IRS demo products were distributed on hybrid pressure levels.

Hybrids levels are suitable for NWP applications but it difficult the normal processing in Nowcasting. Hybrid pressure levels makes difficult the generation of even simple displays or representation of fields. As example the simple display of the temperature at synoptic levels as 500 hPa or 700 hPa is an issue that involves the need of ad hoc functions and processing.

By this reason it was made the interpolation from hybrid pressure levels to a configurable set of fixed pressure levels. This vertical interpolation process introduce some additional advantages:

1. It can be used to reduce the size of the files. As example the specific humidity profiles could be reduced from the 137 hybrid levels to just a configurable set of fixed pressure levels from surface to 200 hPa.
2. The synoptic pressure levels fields could be added. Thus fields as 500 hPa temperature or dew point at 850 hPa are precomputed and available for just display. Advanced display generation as thermodynamical diagrams and vertical cross sections are also simplified with the vertical interpolation to fixed pressure levels.

The interpolation from hybrid pressure levels to a configurable set fixed pressure levels was based in the routine on iSHAI/MSG to interpolate from hybrid levels ECMWF GRIB files to the RTTOV pressure levels.

In the writing of the compatible with McIDAS-V netCDF-3 files the set of configurable fixed pressure levels is added as an additional dimension and the profiles variables are declared to have this extra dimension. Thus, any standard netCDF program or languages as Python can access to pressure level for any array position. The overhead introduced by this pressure level variable is minimal taking into account the benefits.

In the case to maintain the hybrid level representation, the use of one 3D array as temperature will imply that pressure levels should be calculated by users using ad_hoc code and it is more difficult to use external tools to select levels, apply algorithm, creates subset of pressure levels, etc.

5.2 COMPARISON OF FIELDS FROM OTHER SOURCES AND EXAMPLE OF THE DEFICIENCIES ON NRT PROCESSING

In this section it is shown a summary of the comparison of IRS_proxy outputs in the case study of 10th August 2016. On this case study some fast convective development took place in the South of Spain.

Using the process described in Section 5.1 it was converted the NRT IRS-proxy netcdf to McIDAS-V compatible McIDAS-V.

It has been shown and compared here the ML parameter (ML: precipitable water in layer [850-500]) and interactive sounding using the McIDAS-V images from:

a) NRT IRS_PROXY. In the Figure 9 it can be compared the display of ML from NRT IRS_proxy FG (left) and physical retrieval (right). For a better comparison it has been superimposed IR10.8 images on NWCSAF detected clouds and avoid the black hole images comparisons. At the **green** pixel probe of the Figure 9, it has been interactively created the Skew-T diagrams that can be seen in the Figure 12 from NRT IRS_proxy FG (broken line) and physical retrieval (solid line). It is clear a strong change in the IRS_proxy profiles between the FG and physical retrieval step.

b) ECMWF: In the Figure 10 it can be compared the display of ML from ECMWF at 06Z (left) and 10:30Z (right). The ECMWF ML has been generated with the AEMET PGE00 tool that made also the 4D interpolation (vertical, temporal and spatial interpolation) from 0.125º by 0.125º every hour forecast at hybrid levels GRIB files. In the Figure 13 it can be seen Skew-T from ECMWF using 4D iSHAI interpolation at 06:00Z and 10:30Z forecast from the run 00Z of the 10th August 2016 at **green** pixel probe of the Figure 10.

c) NWCSAF iSHAI: In the Figure 11 it can be seen the display of ML from iSHAI at 10:30Z using as input the MSG-SEVIRI IR BTs and the ECWMF 4D interpolated to 10:30Z. In the Figure 14 it can be seen the Skew-T from NWCSAF iSHAI using SEVIRI BTs at **green** pixel probe of the Figure 11.

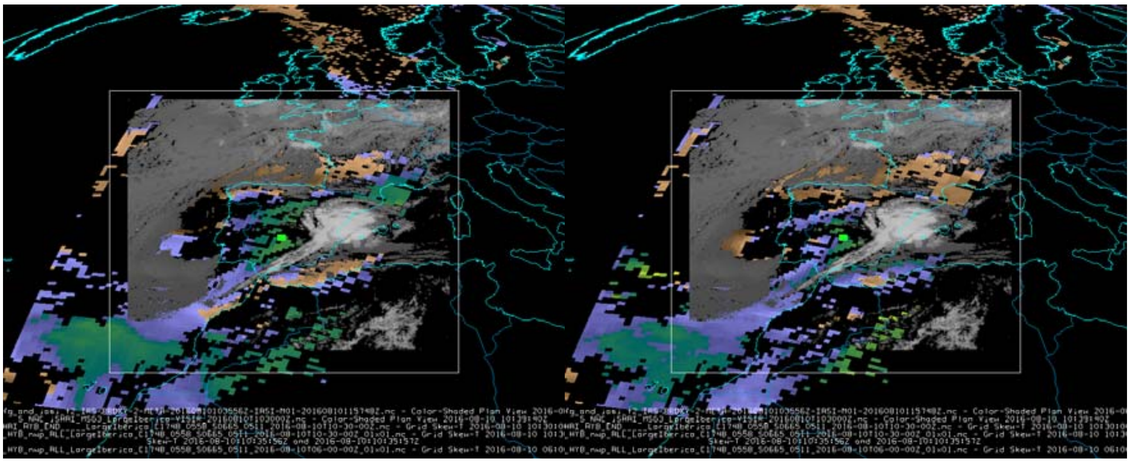


Figure 9: ML parameter from FG IRS_proxy (left) and from physical retrieval FG IRS_proxy (right) at 10:35Z. Overimposed in gray MSG IR10.8 (10:30Z) on cloudy pixels in NWCSAF CMA.

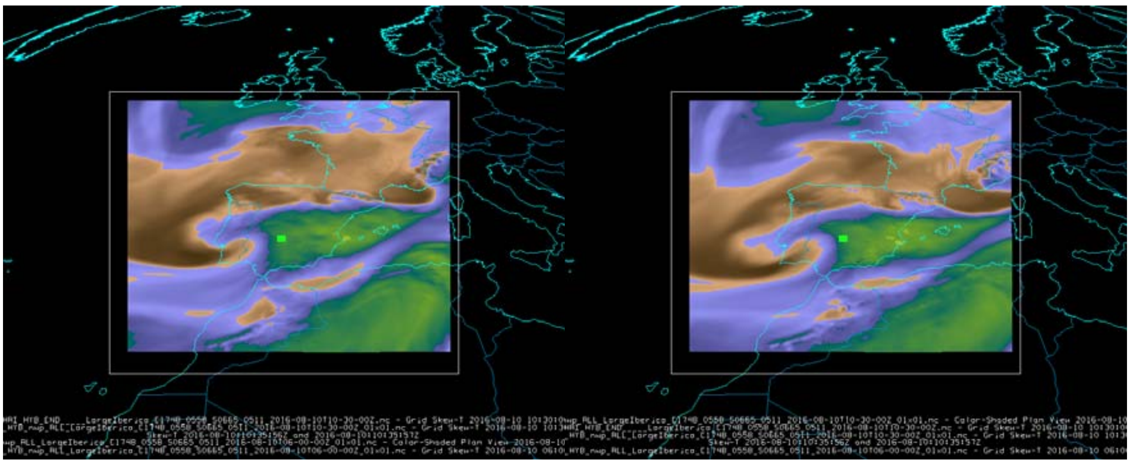


Figure 10: ML parameter from ECMWF t+6:00 (left) and from ECMWF t+10:30 (right)

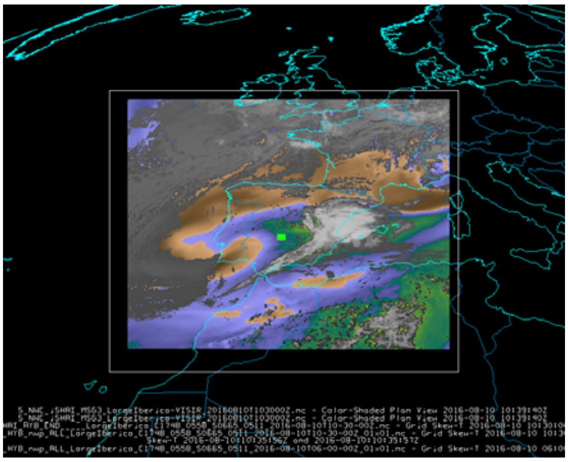


Figure 11: ML parameter from iSHAI 10:30

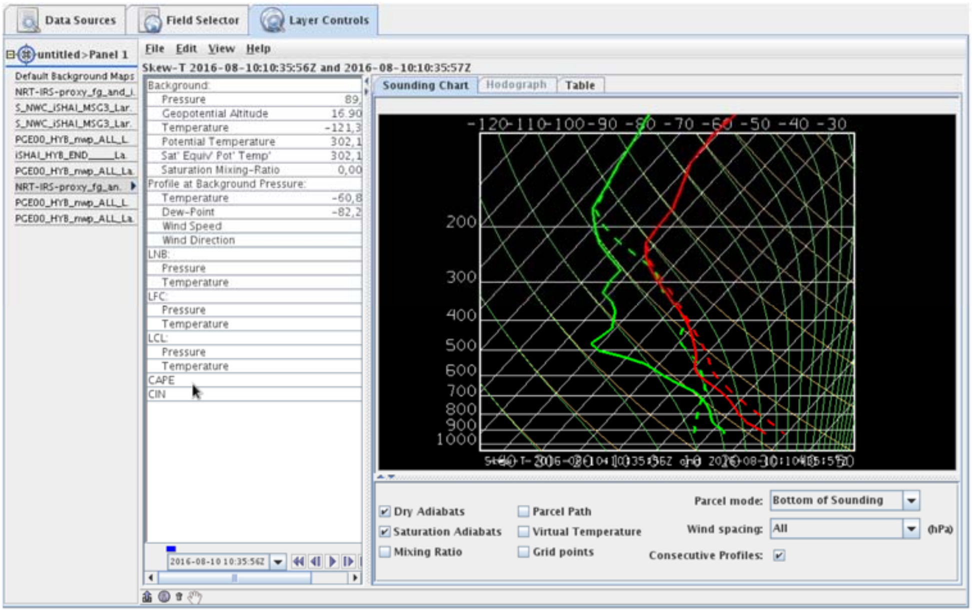


Figure 12: comparison of Skew-T diagrams from IRS_proxy FG (broken line) and from IRS_proxy physical retrieval (solid lines). It can be seen the great change between both profiles.

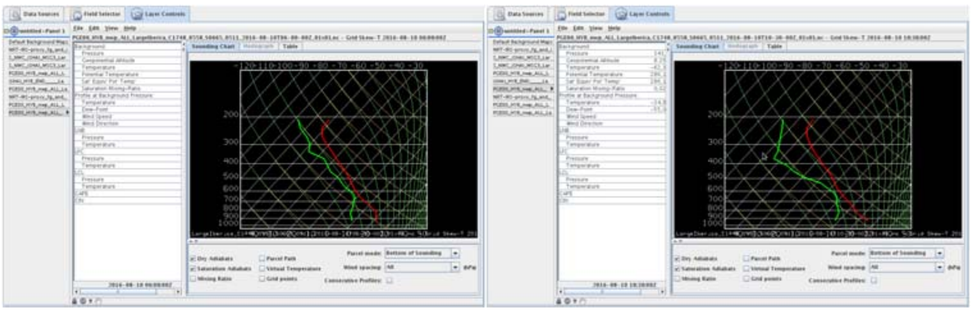


Figure 13: comparison of Skew-T diagram from ECMWF at t+06 (left) and t+10:30 (right).

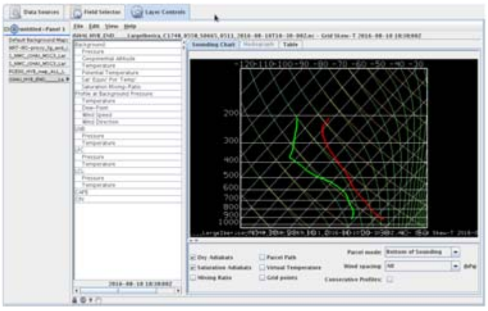


Figure 14: Skew-T diagram from NWCSAF iSHAI from MSG SEVIRI clear air product at 10:30Z.

For a better comparison of the Skew-T diagrams it has been saved the dew point profiles at the probes and plotted together the dew point from the 5 sources.

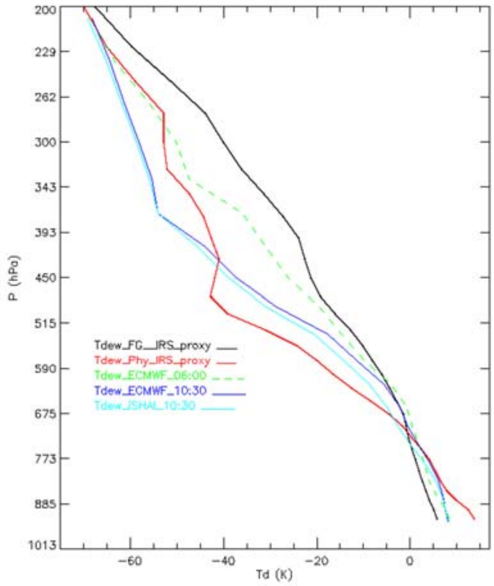


Figure 15: Dew point profile from IRS_proxy, ECMWF and iSHAI profiles at the green pixels on Figures 12 to 14 at 10:30Z.

The comparison of the Figures 9 to 15 allowed to discover the causes of some of the deficiencies in NRT IRS_proxy processing. It is clear in Figure 15 that the FG used (black curve) dew point profile in the NRT IRS_proxy processing is very far to the ECMWF at 10:30Z dew point profile. This likely was due to the use as FG of very aged and from several hours different ECMWF forecast and likely also due to the use of a coarse ECWMF resolution and a not adequate spatial interpolation; it can be seen in Figure 9 (left) there is not present the relative strong spatial gradient in ML ECWMF at 6Z. Since the NRT IRS_proxy FG was wrong and far of the truth, the physical retrieval with real IASI BTs overestimated the correction resulting in the huge difference between FG and physical retrieved profiles in the NRT IRS_proxy processing. The NRT IRS_proxy direction indicates a dryness in the dew point profile. Comparison of ECMWF dew point at 6Z and at 10:30Z profiles confirms the tendency to dryness and agrees with the slight dryness of iSHAI modification to dew point profile of ECWMF at 10:30Z.

To solve part of this deficiency an offline reprocessing was made. The use of offline reprocessed IRS_proxy files is described in Section 5.4.

5.3 COMPARISON WITH THE UMARF EUMETSAT OFFICIAL IASI L2 FILES.

For several cases studies there was experience in downloaded from UMARF the EUMETSAT IASI L2 files; see in References some examples in Martinez [2013, 2015]. On UMARF it is possible to download the IASI L2 files in native or netCDF-4 format. But in any of the cases the UMARF IASI L2 files cannot be used as input to standard display and management tools as McIDAS-V due to lack of adequate design.

The problem of UMARF IASI L2 netCDF-4 is that it has been designed just as transcript from binary to netCDF format without taking into account any rule or convention in order to be compatible with any standard visualization tool. As in UMARF IASI L1 in Section 4.2, this imply that just ad hoc programs or using user developing programs environment as Python can be used to read and use them.

After downloading the UMARF IASI L2 files from EUMETSAT UMARF, it was made the analysis of the format and structure. For an easy use in nowcasting, it was created a GDL and IDL program to convert from the UMARF IASI L2 files files to a McIDAS-V compatible structure netCDF files on fixed pressure levels. The array of fixed pressure levels is customizable by the user (making an interpolation/extrapolation to the array of fixed pressure levels provided in the program). The converter program was developed using as basis previous prototype as the used in Martinez [2013, 2015].

After the experience described in Section 5.1.2.2 it was clear the need to reorder IASI L2 pixels depending on the IASI detectors order. Also the same parameters that in NWC SAF iSHAI product are calculated. In the Figure 16 it is shown an example of McIDAS-V interactive display of IASI L2 version 6.2 after conversion to McIDAS-V compatible netCDF files. It can be seen that the fields are wrong; as example the IASI L2 v6.2 PWLR[3] FG should have not missing values in all pixels. Also that both images have strips of repeated values. The bug in the transcript from native to netCDF was reported and the bug was fixed several months later.

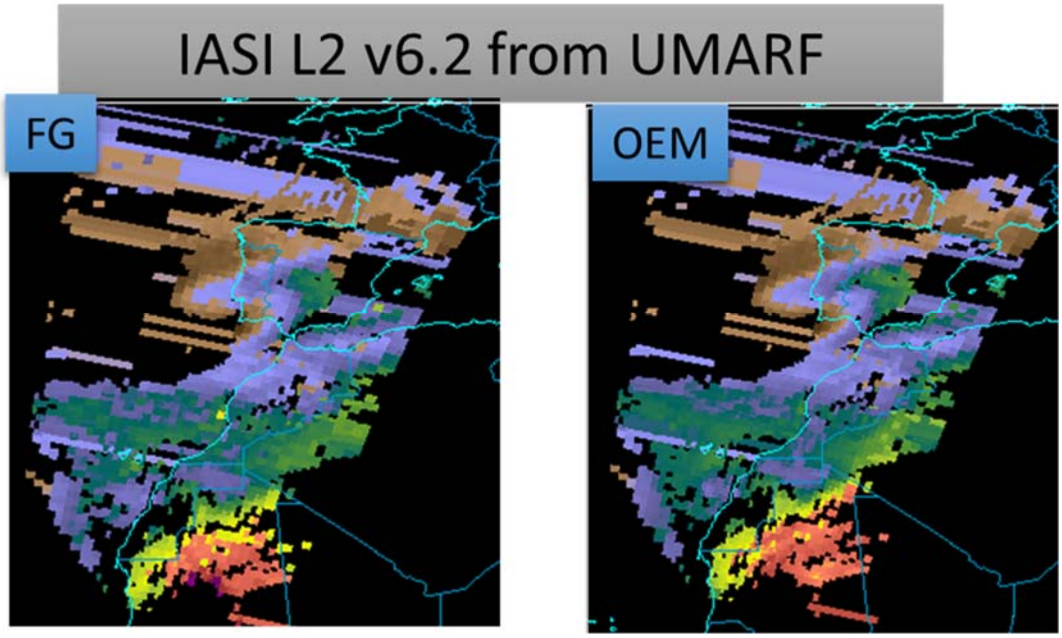


Figure 16: METOPB 20160810 ≈ 10:33Z. From EUMETSAT Secretariat FG and physical retrieval IASI L2 v6.2 download from UMARF the ML parameter has been calculated. (left) IASI L2 v6.2 PWLR[3] FG due to bug in the transcript from native to netCDF not all pixels has value different to missing code and (right) optimal estimation. It was reported the bug and corrected months later.

Just for completeness, in November 2018 inside of CDOP-3 and outside of the reporting period the same IASI L2 were downloaded from UMARF after EUMETSAT fixed the bug in UMARF software. In the Figure 17 it is shown the same fields that in Figure 16 after conversion to McIDAS-V compatible netCDF files after EUMETSAT fixed the transcript of IASI L2 from native to netCDF.

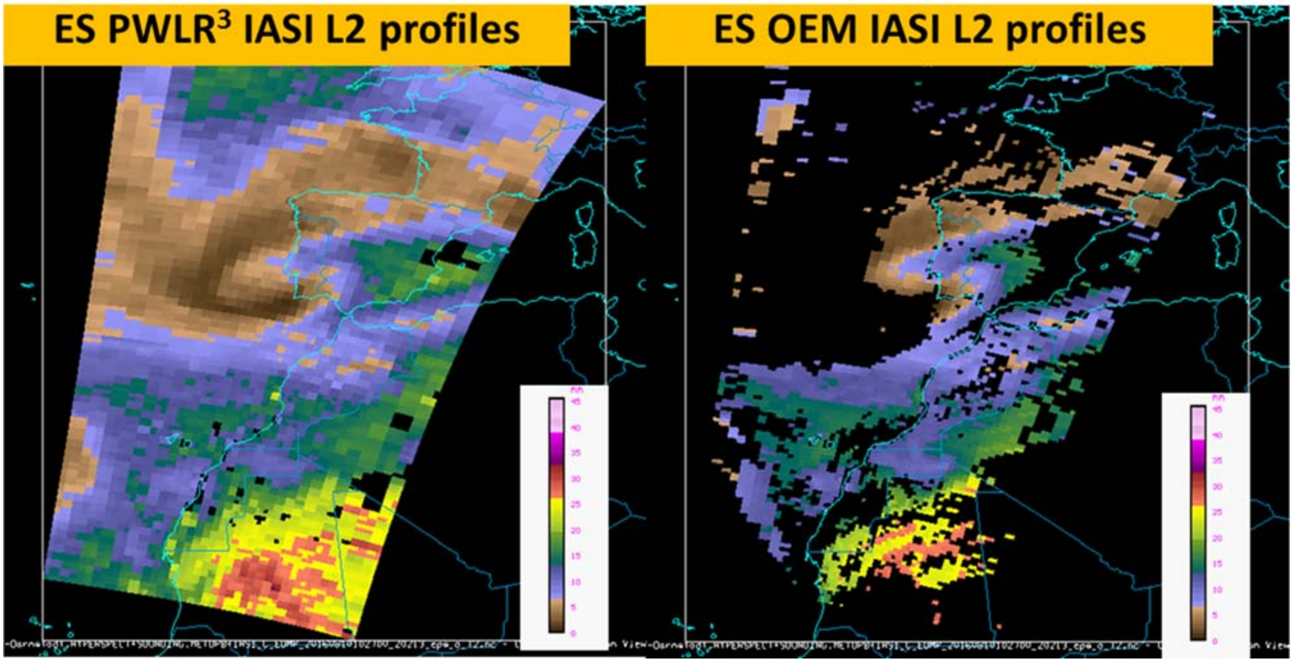


Figure 17: METOPB 20160810 ≈ 10:33Z. From EUMETSAT Secretariat FG and physical retrieval IASI L2 v6.2 ***download from UMARF in CDOP-3 circa November 2018 after EUMETSAT fixed the transcript from native to netCDF****. The ML parameter was calculated. (left) IASI L2 v6.2 PWLR³ FG and (right) optimal estimation.*

This demonstrated that EUMETSAT L2 are written not thinking in to be exploited as image arrays. This is other example of the need of NWC SAF software in the users' side of EUMETCAST to cover the gap between EUMETSAT disseminated files and users' needs. Using McIDAS-V netCDF file this problems are solved and as can be seen in Martinez [2013, 2015 and 2018b] it is possible to generate interactive exploitation of thermodynamical diagrams, 2D display and 3D display (as vertical cross section, etc.).

5.4 COMMENTS AND SUGGESTIONS MADE FOR IMPROVING OF THE MTG-IRS L2 PROCESSING AND NETCDF DESIGN.

A summary of some issues, comments and suggestions made for improving the MTG-IRS L2 processing and the netCDF MTG-IRS L2 was provided (together with an interactive example session) on one of the periodical Webex coordination meetings.

In general the issue of the main comments were related to the concern that after the huge EUMETSAT efforts of IRS L2 products a poor design of the IRS-L2 format could prevent the generalized use by users of MTG-IRS L2 files if they are not compatible with user tools. Many of them are similar to the ones made for MTG-IRS L1 in Section 4.3.

Due to this and the quality of FG issue, it was suggested to explore and investigate on:

- ✓ Need that the IRS-L2 format will follow a standard convention. Similar to the comment for MTG-IRS L1 in Section 4.3.
- ✓ netCDF-4 is a format built over HDF-5 and netCDF-4 are indeed special case of HDF-5 files. Same to the comment for MTG-IRS L1 in Section 4.3.

- ✓ Although the content and structure of example MTG-IRS files may be useful for assimilation into NWP format they have a too complicated structure for nowcasting.

  As commented in Section 5.1.2.3 the representation on hybrid levels is adequate for NWP or some simple use but not for nowcasting.

- ✓ It should be improved the First-Guess supply. It was suggested to follow the way used on the NWC SAF iSHAI 4D interpolation (pressure levels, time, spatial).

  The ECMWF hybrid GRIB files used to generate the First-Guess should have as much spatial and temporal resolution as possible. It was recommended to download from MARS every 1 hour from *t+00* to *t+24* for every run (00 and 12 UTC) with spatial resolution at lest of 0.125ºx0.125º. It should be taken in account that for a global (all the World region) MARS request the amount of data to get and download could generate timeliness problems (MARS could be a bottle-neck for a request as commented above for large regions).

  It was also suggested to avoid the nearest neighbour interpolation and implement a process as the used in the NWC SAF iSHAI using a 4D interpolation. In iSHAI it is made first one interpolation and extrapolation from hybrid levels to a set of fixed pressure levels, second it is made time interpolation from the previous and next GRIB files times and last it is made a bilinear spatial interpolation to the iSHAI pixel.

  Also it was recommend to include one attribute with the date of the ECMWF run and forecast times (example *t+12* or *t+15*) used as First-Guess. This attribute could be used by the users to know the temporal distance and the age of the FG.

  Stephen Tjemkes answered that these recommendations were not possible to implement because also ECWMF forecasted covariance matrix were used in the algorithm and they cannot been made the interpolation process. The temporal resolution was 3 or 6 hours and due to use an ECMWF experimental version (in order to explore the use of near real time forecasted covariance matrices on global coverage) it was created some bottle-neck in the communications between ECWMF and EUMETSAT. This handicaps was the main reason in the use of bad FG in the near real time processing due to use of aged forecast GRIBs.

- ✓ Internal compression. As commented on MTG-IRS L1 in Section 4.3 it is needed to investigate the internal compression.

  In the case of L2 products it should be needed to made tests to write arrays of **FG** and **IASI L2** for all pixels in order *XX (60, nlines, niveles_hibrido)* and *XX (niveles_hibridos, 60, nlines)* and to look what compression rate is better.

  Also was remember the need to use internal compression of netCDF files using standard *gzip* compression. See comments on Section 4.3.

- ✓ The date and time variable must follow also standard practices.

  Also it will be recommend to add additional time dimension to allow concatenation of fields on several netCDF files for one slot in just a netCDF file for a period. Later, using external netCDF tools (as CDO) it will be possible to extract a subset of fields from a set several netCDF files for one slot and combine in netCDF files for a period.

5.5 Reprocessing phase of MTG-IRS NRT Demo Project

To solve the issues of the bad quality of the FG in NRT processing an offline reprocessing was made. In the reprocessing the ECMWF supply to FG input chain runs successfully and MTG-IRS demo reprocessing was better.

Some of the results were presented at the MTG-IRS NRT Demo Project Review Meeting the 18th-19th May 2017 in EUMETSAT. The objective of this workshop was the interchange of experiences with level 1 and level 2 hyper-spectral sounding products in an (ideally) operational weather forecasting environment.

Due to the slow reprocessing rate at the date of May Workshop the 10th August 2016 case study was not reprocessed. Then it was shown at the Workshop the case study of the Malaga flash flood on 19th February 2017 using the McIDAS-V framework. The lack of clear air situation did not allow to extract great conclusions; just to certify the improvement in the IASI fields with the improvement in the FG. It can be seen in Figure 18 the great improvement in the ML field MTG-IRS Demo reprocessing mode and the bad aspect of the fields from the MTG-IRS Demo NRT processing; just the change is due to use more correct FG.

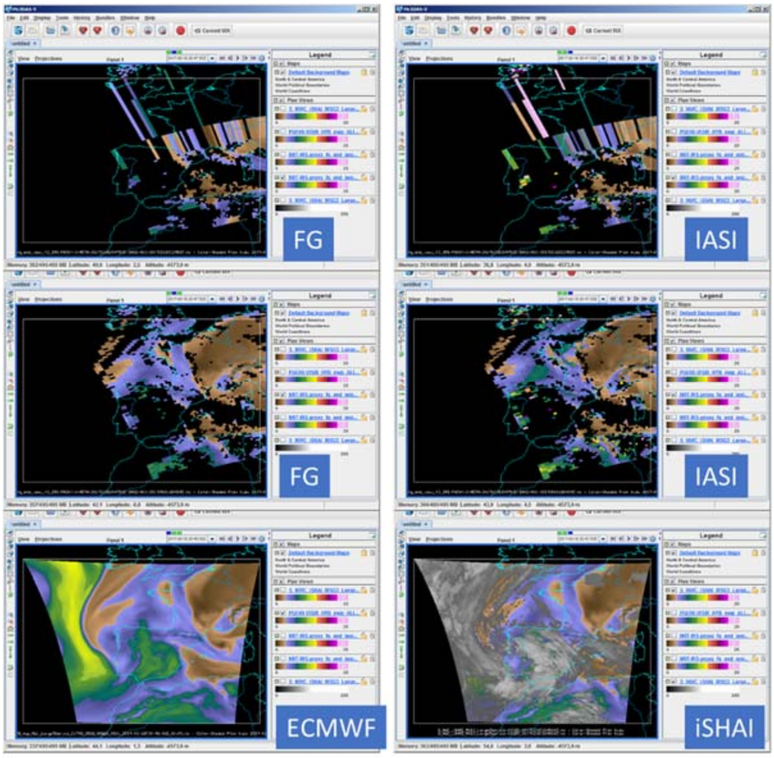
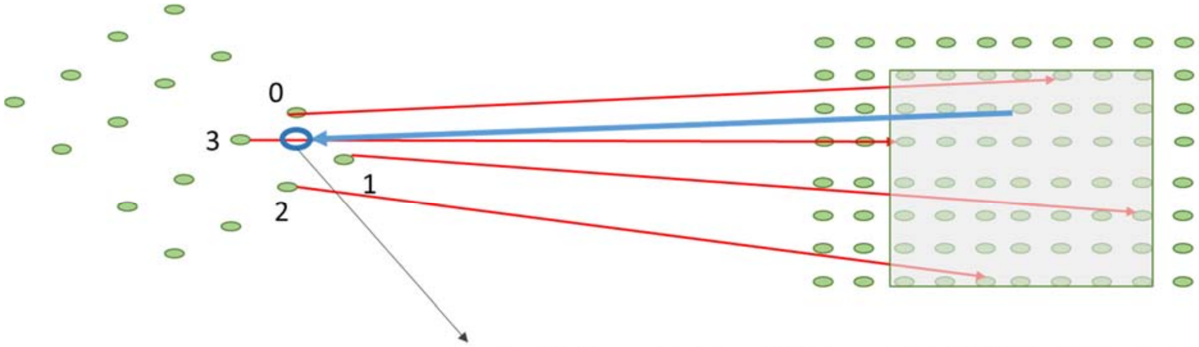


Figure 18: Case study 19th February 2017 at 20:45Z. ML comparison of: (top row) ECMWF FG in NRT proxy (left) IASI physical retrieval in NRT proxy (right). (center row) ECMWF FG in reprocessing (left) IASI physical retrieval in reprocessing (right). (bottom left) ECMWF (bottom right) iSHAI from MSG. ML is the precipitable water in layer [850-500 hPa].

Also in the May 2017 MTG-IRS NRT Demo Project Review Meeting was presented the first version of IDL program to make a fast and correct reprojection from IASI to MSG grid. It was started the study of optimal and fast reprojection algorithm from IASI pixels (in future IRS) to NWC SAF regions (subsets of MSG reprojection grid) to avoid the slow process to calculate for every MSG GRID the nearest pixels to the whole IASI dataset. As can be seen in Figure 19, it is based in:

- ✓ Search valid pixels in IASI array.
- ✓ For every valid IASI pixel use the (longitude, latitude) of the IASI pixels and a function that using as input the position (longitude, latitude) returns the MSG grid positon (column, line). This function is used to calculate the (column, line) in the NWC SAF region for the 4 IASI corners.
- ✓ Then for every MSG pixel between the corners, it is calculated the distances of the MSG pixel to the 4 IASI neighbours and it is calculated the value at the MSG pixel as combination of the value in the 4 IASI neighbours weighted for the inverse of the square of the distances.



$$v= (v_0*f(d_0)+ v_1*f(d_1)+ v_2*f(d_2)+ v_3*f(d_3))/(f(d_0)+f(d_1)+f(d_2)+f(d_3))$$

Figure 19: fast reprojection from IASI to MSG grid.

In the Figure 20 are shown the results of the reprojection algorithm of the IASI L2 product from the reprocessing to a NWC SAF region. It can be seen that Skew-T diagrams at pink probe are similar and there is not the great change in the profiles due to the better quality of the FG.

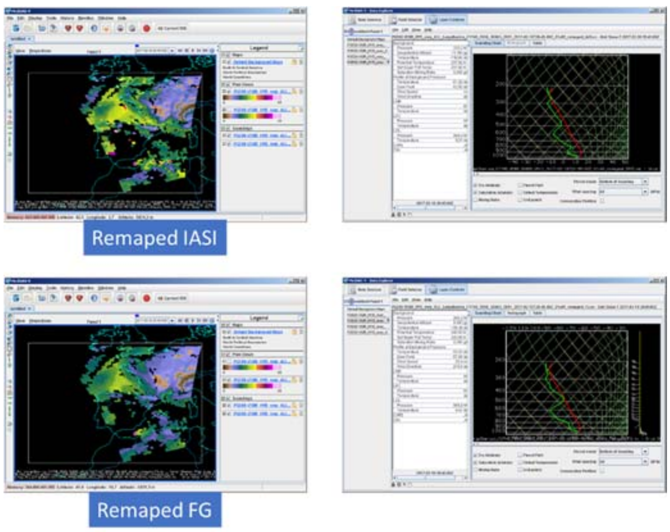


Figure 20: Case study 19th February 2017 at 20:45Z. Reprojected to NWCSAF ML fields and Skew-T diagrams at pink probe. (Upper row) IASI physical retrieval in reprocessing (bottom row) ECWMF FG in reprocessing.

5.6 PROPOSAL OF SSHAI_ES SERVICE.

Due to the limitations commented above the IRS NRT Demo files were never used in an operational environment; in some presentations and discussion the project was shown to some people of the AEMET forecasting department (ATAP) that commented their interest in case to fix them.

The main lesson learned of the activities of MTG-IRS Near Real Time (NRT) Demonstration Project described in this report is that depending on the final Level 2 format specification, the gap between user's tools and MTG-IRS L2 output files could be more or less large **but always some kind of converters tools should be provided to the users**.

The collaboration on the MTG-IRS Near Real Time (NRT) Demonstration Project always has been made as developer thinking how to cover the gap between the downloaded files from EUMETSAT to get them ready to use with tools available in the user side of EUMETCast. The experience allowed to incorporate in the CDOP-3 proposal Nowcasting SAF the need of a service for the support of EUMETAT Secretariat L2 products. This service include the combination and reprojection of several dwells to user defined regions, interpolation to a set of configurable fixed pressure levels, the calculation of derived nowcasting parameters and the writing on netCDF files compatible with users' tools on NWCSAF user defined region of interest.

It was proposed for this service the name ***sSHAI_ES:*** support of the MTG-IRS L2 files disseminated by EUMETSAT Secretariat.

The outputs of the ***sSHAI_ES*** service will be a netCDF file for selected and configurable NWCSAF regions with a combination of 3D and 2D fields. As 3D fields the file will contain the profiles of temperature, specific humidity and ozone on a configurable set of fixed pressure levels at the spatial resolution of MTG-FCI 2x2 km or at the half resolution (4x4 km) FCI grid. As 2D fields the file will contain the skin temperature, surface pressure (provided by the EUMETAT Secretariat L2 products file to reconstruct the hybrid pressure levels), ancillary fields (longitude, latitude, orography, etc) and some parameters of nowcasting interest calculated from the temperature, humidity and ozone profiles. This approach have a lot of advantages:

- ✓ Several MTG-IRS dwell files will be usually needed for covering a user configurable region of interest.
  In the MTG-IRS NRT demo project each file was a granule with 5 minutes of IASI and several IASI L2 file was needed to cover a region.
- ✓ Second the disseminated MTG-IRS L2 dwell files will not be rectified by EUMETSAT to a geostationary common grid.
  In the MTG-IRS NRT demo project each file was on IASI coordinates.
- ✓ The combination and reprojection from IRS grid to user NWC SAF regions on a geostationary grid will be made with the disseminated EUMETSAT Secretariat (ES) MTG-IRS L2 products.
  In the MTG-IRS NRT demo project to get L2 products on a NWCSAF region was needed the process to combine several individual 5 minutes IASI L2 files of one orbit.
  The tasks made are: read several 5 minutes IASI L2 files, reorder the IASI detectors, convert from hybrid levels to a fixed set of configurable pressure levels, calculate the iSHAI NWCSAF parameters (TPW, LPW and stability indices, relative humidity profile, etc), remap the 2D and 3D variables to MSG grid and write on McIDAS-V compatible netCDF-3 format.
  Then, 2D and 3D fields read from the IRS L2 dwells files will be combined and reprojected to users NWC SAF regions.
- ✓ Most of the nowcasting users can truncate the vertical extension of the profiles to the used ones.
- ✓ Calculation of several nowcasting parameters (TPW, LPW and instability indices) will be made also if needed.
  EUMETSAT idea has changed with time from no parameters calculation to some; but in any case due to bandwidth constraint just a few could be disseminated. But in users' side the software could be update to calculate all the parameters needed by users.
- ✓ The idea is that ES MTG-IRS L2 fields could be used directly by users in combination with the ones from MTG-FCI fields for same regions. As example comparison of MTG-FCI/MTG-IRS/NWP instability indices.

- ✓ The fields remapped to NWCSAF regions could be used as input to local generation for NWCSAF MTG-FCI or MTG-IRS products. Field as IRS L2 SKT or IRS L2 emissivity fields could be very useful to improve the MTG-FCI product in future versions. This way a synergetic use of the MTG mission could be get.

6 CONCLUSIONS, SUMMARY AND FUTURE ACTIVITIES

Validation and training dataset generation is a continuous and important task. The EUMETSAT MTG-IRS NRT Demonstration Project did not provide an adequate validation and training dataset with IASI but it provided the tools to build it.

It was foreseen a second edition of the EUMETSAT MTG-IRS NRT Demonstration Project. But due to the retirement of the at that time MTG program responsible Rolf Stuhlmann, the second edition was cancelled on January 2018. Since the MTG-IRS NRT Demonstration Project was stopped and the proposed modifications to the EUMETSAT processor (e.g. better description of the First Guess, improvement of the physical retrieval algorithm, improved netCDF design, etc.) was not implemented, the migration of IDL to GDL (to be used freely by other users) of the whole prototype was not completed. For this reason the idea of the migration of IDL prototype to a freely distributed GDL prototype was also stopped.

Some of the IDL/GDL functions of the prototypes developed in this activities will be needed in CDOP-4. One option is the writing of full GDL prototype; the other options are the developing of the prototypes on Python or directly in C or Fortran-90. The main problem for a direct migration from IDL/GDL to Python is that the algorithm of developed IDL/GDL prototypes uses structures arrays and some IDL/GDL commands that needs pointers. The decision of the migration from IDL to GDL or Python or Fortran/C as proxy or future NWCSAF developments will be taken later during CDOP-3 or early CDOP-4 phase.

The works, studies and ideas described in this report will help to implement NWC SAF use and generation of NWC SAF products and services for MTG-IRS L1 and L2. These studies are also valuable to anticipate the new ways that MTG-IRS L1 data and L2 products could be used in MTG-IRS era for nowcasting by users.

At the same time these activities were made on the NWC SAF framework, the AEMET NWCSAF team was writing at this time the proposal for CDOP-3 with the inclusion of the services in order nowcasting users' community could get a full synergy and provision of imagery and products from MTG/FCI and from MTG/IRS.

So, to summarize, the NWC SAF will deliver an optional and complementary software package in order nowcasting users could make use of MTG/IRS L1 data disseminated in PC, together with EUMETSAT Secretariat IRS L2 products and locally generated NWCSAF MTG-IRS L2.

The NWC SAF is developing computer packages that will allow to manage the data from all the three MTG instruments and to generate locally products by the users. Since NWC SAF provide software that it is installed on the users' computers there are not bandwidth constraints and the whole set could be used in combination with local NWP and data. Thus, all the synergies can be exploited.

To say that a file is written in netCDF only means that it is a needed the netCDF of HDF-5 libraries to open and read it. It is also needed to take care in the design of the internal structure of the netCDF file. The details in the design of the internal structure of a netCDF files are very important; just a small bug in a declaration, dimension or attribute of a variable is the different of netCDF file be usable by several meteorological applications or not. The design of the internal structure of a netCDF files should be as closer and as flexible to the user tools as possible. The optimal format be one that will allow the users just "click and play" files and some coordination should be promote in the meteorological and satellite community.

7 REFERENCES

Martinez M. A.; 2020. Preparation of MTG era: developing of imager and sounder nowcasting products and services for pre-convection phase. NWCSAF Users' Workshop 2020. *Link to* pptx *in NWCSAF web page. Include also links to case studies web page and loops in NWC SAF web page.*

Martinez M. A.; Calbet X, 2019. iSHAI and PGE00: key tools for preconvective monitoring and for the preparation of the MTG era. *EUMETRAIN Convection Event Week 2019. Link to pptx in NWCSAF web page. Link to recording and PDF in EUMETRAIN web page with HAnis loops in NWC SAF web page.*

Martinez M A.; 2018b. Action.M5.A14 - NWC-SAF services and products for MTG-IRS. *IRS Mission Advisory group (IRS-MAG) November 2018 in Brussels, Belgium. Link to PDF in EUMETSAT web page with HAnis loops in NWC SAF web page.*

Martinez M. A.; Calbet X, 2018a. iSHAI and PGE00 as key tools in pre-convection. 2018 *Convection Working Group Workshop in Ljubljana, Slovenia. Link to pptx in CWG web page. Link to PDF in CWG web page with HAnis loops in NWC SAF web page.*

Martinez M.A.,; Calbet X, 2016b. Instability indices and water content in clear air NWCSAF product. *2016 Convection Working Group Workshop in Florence, Italy.*

Martinez M.A.,; Calbet X, 2016c. NWCSAF/GEO clear air product version 2016. Evolution to the MTG era. *Proc. The 2016 EUMETSAT Meteorological Satellite Data User's Conference, Darmstadt, Germany. (S2_Martinez_Miguel.pdf) in zip with oral proceedings in Nowcasting Session. Zip*

Martinez M.A.; 2016a. Studies for comparison of NWCSAF/MSG PGE13 SPhR and IASI L2 products. *NWC/CDOP2/GEO/AEMET/SCI/RP03, Issue 1, Rev.0 28-January-2016. Available in NWCSAF Help Desk private document archive.*

Martinez 2015. Presentation in the 2015 NWC SAF Workshop of the former PGE13 SPhR. See the loops and several examples of the use of vertical cross sections for nowcasting for case studies 12th August 2011 and 20th June 2013. Also the ideas for generalization to other GEO satellites and the extension to MTG-IRS applications in nowcasting are shown.

Martinez M.A., Calbet X., 2013. *Innovative ideas for using the hyperespectral level 1 data of the next geostationary MTG-IRS in nowcasting.* Proc. of the 2013 EUMETSAT Conference, Vienna, Austria.

Martínez M.A., X. Calbet, Prieto J., Tjemkes S. 2010. *Use of synthetic RGB images in training.* Proceedings of the 2010 EUMETSAT Meteorological Satellite Conference, Cordoba, Spain.

ANNEX I: NWCSAF AEMET´S PROPOSAL FOR THE "DEMONSTRATION OF NEAR REAL TIME SERVICES OF MTG-IRS PROXY DATA"

NWCSAF AEMET´s Proposal for the "Demonstration of Near Real Time services of MTG-IRS proxy data"

At AEMET we believe that the MTG-IRS instrument has a very strong potential for Nowcasting. The NWCSAF is currently providing several Nowcasting products to its users based on data from geostationary satellites such as Meteosat (MSG). The natural evolution of some of these products is to use the added capabilities of MTG-IRS, which provides an enormous increase on channels with respect to its predecessors (e.g. SEVIRI), as inputs to the future products that will be provided by the NWCSAF. As such we are interested in the "Demonstration of Near Real Time services of MTG-IRS proxy data" project.

In line with our current and future portfolio of Nowcasting products we can contribute with several concepts that we believe will be very useful for Nowcasters. Before detailing which concepts we can contribute to, it is worth noting some issues with the current proposal:

1. It should be noted that currently we do not have access to the GEANT network. As an alternative, some files could be provided on some ftp site where we could fetch them. For the sole purpose of demonstrating the NWCSAF products tailored for MTG-IRS it is not necessary to have a stream of near real time products nor is it necessary to process all MTG-IRS data. Therefore, providing a small sample of products on an ftp site would be sufficient. The geographical area that could be served on the ftp site could be limited to, as an example, a square centered in Madrid and extending out around 1500 km.

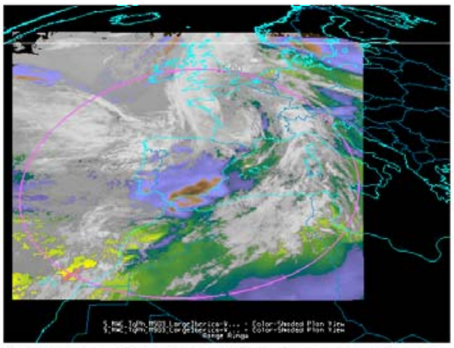


Figure 1: Example NWCSAF TPW display on region interest .

2. The "Scaled Projected States" and their associated error covariance matrices would not be needed by us.
3. Several of the NWCSAF products have been designed for MTG-IRS L1 proxy products. It would be beneficial to have available the MTG-IRS L1 proxy products which correspond to the MTG-ITS L2 ones. This would be beneficial to combine all products in synergy and verify its potential use for Nowcasting.

The concepts and products the Nowcasting SAF can contribute with are:

1. Analyze in detail the format of the MTG-IRS L1 proxy and L2 data such that it is useful for Nowcasters. The MTG-IRS data supplied should have an officially supported format, i.e., one similar to the one that will be later used operationally for dissemination. We

could provide feedback to EUMETSAT on the properties of the format such that it is useful for Nowcasters. Also, to display this data conveniently the MTG-IRS L1 proxy image would need to be overlaid to cover pixels that are cloudy. To verify the usefulness of this approach, Figure 1, which has overlaid L1 data, can be compared to the Figure 2 which does not have any overlaid L1 data. This fact is very important for Nowcasting in loops and in cases where near all pixels are cloudy.

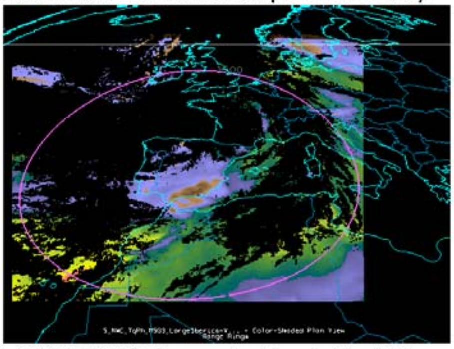


Figure 2: Example NWCSAF TPW display on region interest without IR image on cloudy pixels.

2. If not calculated at EUMETSAT Secretariat, calculate atmospheric stability parameters and water vapor content from the MTG-IRS L2 profiles.
3. Several MTG-IRS granules could be geolocated, assembled and re-projected to form a uniform field, without any discontinuities, over the area of interest specified by the Nowcaster user. Using McIDAS as the foreseeable tool for the NWCSAF MTG-IRS, we will identify the key points to achieve this and perform tests with parameters of interest in 2D (TPW, LPW and stability index).
4. ECMWF fields could also be used to calculate instability parameters and water vapor content. The fields themselves and these mentioned parameters could then be compared with the ones provided by MTG-IRS L2. In the PowerPoint presentation from Martinez (2015), you can verify the convenience to compare interactively with ECNMWF fields (see as an example slides 12 to 21 from Martinez, 2015).
   These comparisons are made with tools available by the NWCSAF but which are generally not distributed to NWCSAF users and are usually used for validation and dataset generation. The differences between MTG IRS and collocated ECMWF will be shown. It will also be investigated the difference between analisys t+00 and t+12.
5. Other derived interested 3D fields is the equivalent potential temperatures, as has been shown by Ralph Petersen from CIMSS-Wisconsin. These 3D fields could be obtained from MTG-IRS like profiles and ECMWF (see slides 24 to 25 from Martinez, 2015).
6. Several Meteosat like RGBs, obtained from the MTG-IRS L1 proxy data, could also be displayed and compared to the other products. In we do not have access to the MTG-IRS L1 data, one IR image from MSG could be used. Also, a comparison with NWCSAF cloud mask could be subjectively done.
7. All sky retrievals could be obtained from the MTG-IRS L1 proxy data. These, and any derived products from them, could then be compared to the rest of the dataset (ECMWF, MTG-IRS L2).

References

Martínez, M.A., 2015, Presentation at the Nowcasting User Workshop, http://www.nwcsaf.org/WorkshopsTrainingSurveys/2015UsersWorkshop/2015UsersWSPresentations/SESSION_III/martinez_nwcsaf_workshop_madridfeb2015.pps

ANNEX II: TECHNICAL NOTE: COMMENTS ON THE STRUCTURE OF SAMPLE 2016 IRS L1 NETCDF FILES.

The EUMETSAT Network of Satellite Application Facilities

NWC SAF

Support to Nowcasting and Very Short Range Forecasting

Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files

NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0, Issue 1, Rev.0

8 November 2017

Based on tests with sample 2016 IRS L1 files for Topical Meeting on Level 1 Processing and IRS L2 files from NRT-IRS proxy experiment.

Prepared by Miguel A. Martinez (AEMET and NWC SAF)

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	**Code:** NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 **Issue:** 1.0 **Date:** 8 November 2017 **File:** Technical_note_comments_IRS_netcdf_structure.docx **Page:** 2/28

REPORT SIGNATURE TABLE

Function	Name	Signature	Date
Prepared by	Miguel A. Martínez (AEMET)		*8 November 2017*
Reviewed by	Xavier Calbet (AEMET)		*8 November 2017*
Authorised by	Xavier Calbet (AEMET)		*8 November 2017*

DOCUMENT CHANGE RECORD

Version	Date	Pages	Changes
1.0	*8 November 2017*	28	

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	**Code:** NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 **Issue:** 1.0 **Date:** 8 November 2017 **File:** Technical_note_comments_IRS_netcdf_structure.docx **Page:** 4/28

Table of Contents

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	Code: NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 Issue: 1.0 Date: 8 November 2017 File: Technical_note_comments_IRS_netcdf_structure.docx Page: 5/28

List of Tables and Figures

1 RATIONALE AND INTRODUCTION

In February 2016, EUMETSAT organized a Topical Meeting on MTG-IRS Level 1 Processing with experts, the science working group (SWG) and the MTG-IRS Mission Advisory Group; the main objective of the Meeting was to advance on IRS L1 processing.

As one of the preparatory activities for this meeting, EUMETSAT distributed a sample netCDF IRS L1 file for a dwell in order to familiarize with the IRS L1 format and to make comments on them. This sample file contains a dwell simulation with synthetic IRS spectra from NWP profiles on the area of the Alps and the Gulf of Genoa (also it was distributed in principal components representation on other netCDF file); the area covered can be seen in Figure 1.

Using files in the McIDAS-V tool is a practical way to test their formats. If these files work well with McIDAS-V, it is very likely they will be easy to use with any other tool. I tried unsuccessfully to use it with the McIDAS-V tool (using McIDAS-V as touchstone of future user's tools). From IASI past tests, see Martinez (2013b, 2015, 2016), I have converter IDL programs from EUMETSAT UMARF IASI L1 and L2 netCDF files to McIDAS-V compatible netCDF files. Then, I made other IDL converter program to read the IRS L1 spectra in the sample 2016 IRS L1 netCDF files and write them on a McIDAS-V compatible netCDF file. See Figure 1.

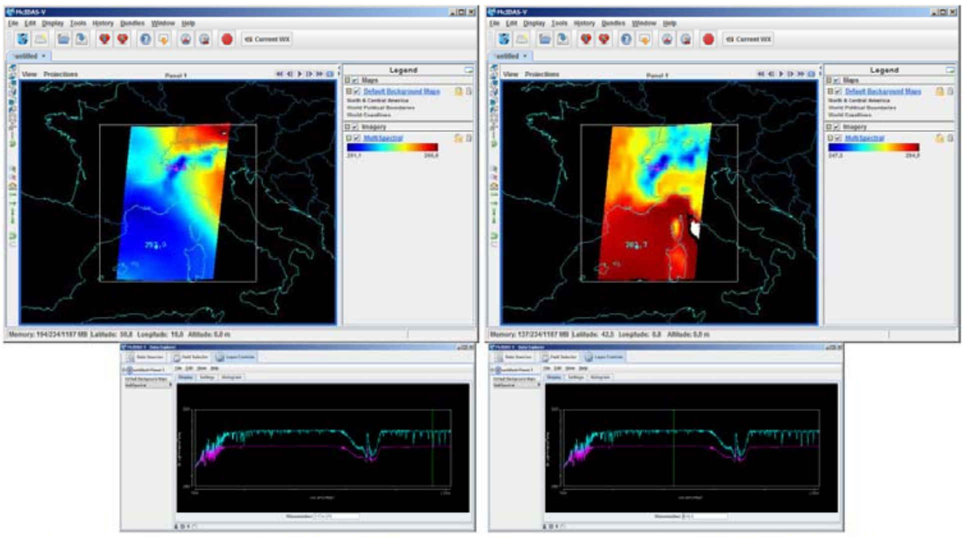


Figure 1: Example of McIDAS-V interactive displays from sample 2016 IRS L1 dwell after conversion of EUMETSAT sample 2016 IRS L1 netCDF files to McIDAS-V compatible netCDF files. (top) brightness temperature at two wavenumbers. (bottom) show the spectra on the pink and blue pixels in the top images on LWIR MTG-IRS band. Users can use move interactively the green lines on bottom images to change the wavenumbers of images in the top; moving the pink and blue cross compare the spectra at different pixels.

The main comments on this document are based in a personal log that describes the main differences between the structures in both netCDF files. It was prepared before the IRS Topical meeting in February 2016 and it was commented in private talks with some EUMETSAT people.

The main recommendation of the Topical IRS Topical Meeting in February 2016 on the IRS format was to discuss deeper in the Ljubljana MTG-IRS Mission Advisory Group Workshop in May 2016. At Ljubljana, I commented and raised some deficiencies in the EUMETSAT proposed netCDF structure that in my opinion could compromise the final objective to get a design of L1 and L2 files adequate for **both** EUMETSAT and user's purposes. After my comments, the IRS MAG 3 stablished the action **IRSMAG.M3.A.2.**

> *"NetCDF file format is chosen as the baseline to prepare the level 1b files. M. Martinez argued that this file format has limitations, which will affect the users. To identify the limitations of the NetCDF file format for operational applications, M. Martinez agreed to report on non-compliances and will propose potential solutions at the next MTG-IRS MAG meeting (action **IRSMAG.M3.A.2**)".*

This technical note has been made in response to this action **IRSMAG.M3.A.2.**

This document is provided also in the framework of the SAF on Support to Nowcasting and Very Short Range Forecasting, hereafter NWC SAF. The main objective of NWC SAF is to provide, further develop and maintain software packages to be used for Nowcasting applications of operational meteorological satellite data by National Meteorological Services.

NWC SAF has the compromise in CDOP-3 to develop prototypes for MTG-IRS services. One of the IRS services is the use in real time of netCDF IRS L1 files in nowcasting. Thus, the search of an adequate netCDF structure is of great importance also for the NWC SAF. It would be desirable that the netCDF files from EUMETSAT Secretariat and NWC SAF would have the same or similar structures and that they could be used by the same tools.

For this reason, this document should be understood also as an early preparation document of the NWC SAF MTG/IRS processing package to exploit MTG/IRS synergistically with the NWC SAF processing package for image geostationary meteorological satellites, NWC/GEO.

1.1 PURPOSE AND SCOPE OF THE DOCUMENT

The structure of the sample 2016 IRS L1 netCDF files proposed by EUMETSAT is suitable for use only with specific and ad hoc tools or using specific and tailored programs. This implies to limit the use for simple visualizations (representation in a map just plotting on (x,y) coordinates), ingest in national meteorological systems and the use in NWP. In other words, repeat the way in which the IASI L1 files are currently manipulated that has led to under exploitation because it requires special programs for reading and data utilization.

Following the approach proposed in this technical note, the files could be used not only with the specific previous mentioned tools but they could also be used with other general purpose tools. This would facilitate the use by many users and also will allow to combine synergistically the IRS data and IRS products with other data sources and NWP. It would also facilitate the processing and manipulation of data by users (even by uninitiated users). Finally, it would also facilitate the use of non-traditional programming languages such as Python.

In section 2, the detected issues on the sample 2016 IRS L1 netCDF files are described and the way they were solved in the McIDAS-V compatible version are shown. Also more ideas for future implementation are commented.

In section 3, a summary with the proposed structure for IRS netCDF files and final recommendations are made.

1.2 DEFINITIONS, ACRONYMS AND ABBREVIATIONS

AEMET	Agencia Estatal de Meteorología Meteorology State Agency (Spain)
ATBD	Algorithm Theoretical Basis Document
BL	Precipitable water in low layer (P_{sfc} – 850 hPa)
BT	Brightness Temperature
CDOP (CDOP-1)	Continuous Development and Operations Phase (1)
CDOP-2	Continuous Development and Operations Phase 2
CDOP-3	Continuous Development and Operations Phase 3
CF	NetCDF Climate and Forecast (CF) Metadata Conventions
CPU	Central Processor Unit
DEM	Digital Elevation Model
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCI	Flexible Combined Imager (MTG)
FG	First Guess
FOV	Field Of View
FOR	Field Of Regard
GEO	Geostationary Satellites
GEO-iSHAI	GEO imaging Satellite Humidity And Instability
GRIB	Gridded Information in Binary Form
HDF5	Hierarchical Data format version 5
HL	Precipitable water in High Layer (500 – 0 hPa)
hPa	Hecto Pascal
IDL	Interactive Data Language
IR	InfraRed
IRS	Infrared Sounder (MTG)
iSHAI	imaging Satellite Humidity And Instability
MARS	ECMWF Meteorological Archive and Retrieval Facility

McIDAS	Man Computer Interactive Data Access System
ML	Precipitable water in Medium Layer (850 – 500 hPa)
MSG	Meteosat Second Generation
MTG	Meteosat Third Generation
MTG-FCI	Meteosat Third Generation Flexible Combined Imager
MTG-IRS	Meteosat Third Generation Infra Red Sounder
netCDF	Network Common Data Form
NRT	Near Real Time
NWC	Nowcasting
NWC/GEO	Geostationary part of the Nowcasting SAF
NWCSAF	Nowcasting SAF
NWP	Numerical Weather Prediction
NWP SAF	SAF for Numerical Weather Prediction
LPW	Layer Precipitable Water
RTM	Radiative Transfer Model
RTTOV	Radiative Transfer for TOVs
SAF	Satellite Application Facility
SEVIRI	Spinning Enhanced Visible InfraRed Imager
SHAI	Satellite Humidity And Instability
SKT	Skin Temperature
SW	Software
TPW	Total Precipitable Water
UM	User Manual
VR	Validation Report
VSA	Visiting Scientist Activities
WV	Water Vapour Channel

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	**Code:** NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 **Issue:** 1.0 **Date:** 8 November 2017 **File:** Technical_note_comments_IRS_netcdf_structure.docx **Page:** 10/28

1.3 REFERENCES

1.3.1 Applicable Documents

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority. They are referenced in this document in the form [AD.X]

For versioned references, subsequent amendments to, or revisions of, any of these publications do not apply. For unversioned references, the current edition of the document referred applies.

Current documentation can be found at the NWC SAF Helpdesk web: http://www.nwcsaf.org.

Ref.	Title	Code	Vers
[AD.1]	Proposal for the Third Continuous Development and Operations Phase (CDOP-3)		1.0

Table 1: List of Applicable Documents

1.3.2 Reference Documents

The reference documents contain useful information related to the subject of the project. These reference documents complement the applicable ones, and can be looked up to enhance the information included in this document if it is desired. They are referenced in this document in the form [RD.X]

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

Ref.	Title	Code	Vers	Date
[RD.1]	Data Output Format for the NWC/GEO	NWC/CDOP2/AEMET/SW/DOF	1.1	11/04/16
[RD.2]	User Manual for the Clear Air Product Processors of the NWC/GEO: Science Part	NWC/CDOP2/GEO/AEMET/SCI/UM/ClearAir	1.0	15/10/16

Table 2: List of Referenced Documents

2. COMMENTS ON SAMPLE IRS NETCDF STRUCTURE

The netCDF format has been chosen for MTG-IRS L1 and L2 files. In principle this only means that to open, manipulate and write the IRS data and products a tool that can handle the netCDF (or HDF-5) library will be needed.

But the netCDF and HDF-5 formats are general-purpose formats that allow data to be stored with an internal structure. The internal structure of one netCDF file can be designed with great flexibility (so that these formats can be used to store any type of data from biomedical data to galaxies).

Therefore, the next step is to adequately define an internal structure. Minor issues such as the name of the variables, the dimensions, the type of data used, etc. can be customized and must be fixed by the user. This apparently naive step has a lot of posterior implications.

At this point in time, there is no a general rule to establish a common structure to be used for similar data. As an ordination attempt, several **conventions** have been created. One of them is the CF (Climate and Forecast Metadata) convention. But normally this conventions just establish one standard in how to name variables for the data fields and the units. See Annex A.1.

The IRS L1 and L2 netCDF files should be:

Easy to use: design the internal netCDF structure to be ready to use for standard tools.

One key concept in netCDF files is the concept of **dimension**. **Dimension** it is a two side concept; in one hand is used to fix the sizes of elements of one array to store a data field, and in the other hand it is an associated data field that with the same size contains the real values. As example, since pressure levels are usually float values and data arrays indices are integer the dimension concept allows stablish a link between one index or position in one array and the real value associated at this position in the array. One of the main advantage of netCDF is the concept of dimensions.

The structure is dwell based with primary time, longitude and latitude dimensions.

Standard conformant (CF) with the objective to be able to exchange and make use of information with actual and future computer languages (C, FORTRAN, Python, Java, etc.) and tools.

- The CF (Climate and Forecast Metadata) objective include:
  - Locate data in space–time and as a function of other independent variables, to facilitate processing and graphics. It is made using or declaring dimensions.
  - Identify data sufficiently to enable users of data from different sources to decide what is comparable, and to distinguish variables in archives.
  - Conventions for dimensions, units, and variable and file attributes.
  - Users will need to consult an interface specification document much less

To use a common structure with similar products from other instruments, the files should be similar but flexible enough to incorporate the differences needed. This implies the need of a common framework for EUMETSAT and perhaps CGMS; it could be made through a satellite convention.

The IRS files should be read and usable by as much tools as possible. Example tools: McIDAS-V, IDV, CDO, Python, Panoply, etc.

NWC SAF AEMet	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	Code: NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 Issue: 1.0 Date: 8 November 2017 File: Technical_note_comments_IRS_netcdf_structure.docx Page: 12/28

netCDF4 allows a hierarchical structure – "groups". But too much use of the group structure imposes a burden on users, so we keep it simple IRS L1 and L2 files; should maintain all of the most commonly used variables in the root group and not produce deep group trees. See 2.8 discussion.

2.1 GEOREFERENCING ISSUE

It is needed to establish correctly the georeferencing attributes, to define correctly variables and expect that as many tools as possible could understand them.

Making a comparison of the design and variables of sample 2016 IRS L1 netCDF files versus McIDAS-V compatible files and NWCSAF version 2016 files, I think that there is a problem in the attributes of some variables in the sample 2016 IRS L1 netCDF files.

In the sample IRS L1 netCDF files the *latitude* and *longitude* are correctly declared and written but they are not correctly associated to the variables. The *latitude* and *longitude* are associated to the georeferenced variables using the wrong attribute *ancillary_variables* and the attribute **coordinates is not present**. In other words, it is recommended to have the **coordinates** attribute in every variable (radiances, for example) **and it** should be associated to the latitude and longitude fields.

```
short effective_radiance(dwell, dwell_row, dwell_column, spectral_sample) ;
        long_name = "Effective radiance" ;
        standard_name = "effective radiance in wavenumber" ;
        ancillary_variables = "spatial_sample_quality radiance_scale_factor radiance_offset latitude longitude" ;

float latitude(dwell, dwell_row, dwell_column) ;
        units = "degrees_north" ;
        long_name = "Geolocation latitude for spatial sample" ;
        standard_name = "latitude" ;

float longitude(dwell, dwell_row, dwell_column) ;
        units = "degrees_east" ;
        long_name = "Geolocation longitude for spatial sample" ;
        standard_name = "longitude" ;
```

Table 3: ncdump of effective_radiance, longitude and latitude fields in the sample IRS L1b netCDF files.

From IASI past tests, see Martinez (2013b, 2015, 2016), I have converter IDL programs from EUMETSAT UMARF IASI L1 and L2 netCDF files to McIDAS-V compatible netCDF files. Then, I made other IDL converter program to read the IRS L1 spectra in the sample 2016 IRS L1 netCDF files and write them on a McIDAS-V compatible netCDF file (see Figure 1). In the McIDAS-V compatible version of IRS L1b netCDF files and in the NWCSAF version 2016 netCDF, the longitude and latitude are named variables (*lon*, *lat*) and they are associated **to every variable using the attribute coordinates**. At least for the McIDAS-V tool, this way to specify the attributes for the georeferencing works.

```
ushort ishai_tpw(ny, nx) ;
       standard_name = "ishai_tpw" ;
       long name = "iSHAI Total Precipitable Water" ;
       …
       ancillary_variables = "ishai_status_flag ishai_conditions ishai_quality ishai_tpw_pal" ;
       coordinates = "lon lat" ;
       _FillValue = -1US ;
```

```
float lon(ny, nx) ;
         CoordinateAxisType = "Lon" ;
        standard_name = "longitude" ;
        long_name = "Longitude at the centre of each pixel" ;
        units = "degrees_east" ;
        valid_range = -180.f, 180.f ;
         FillValue = -9999.f ;

float lat(ny, nx) ;
        _CoordinateAxisType = "Lat" ;
        standard_name = "latitude" ;
        long_name = "Latitude at the centre of each pixel" ;
        units = "degrees_north" ;
        valid_range = -90.f, 90.f ;
        _FillValue = -9999.f ;
```

Table 4: ncdump of effective_radiance, longitude and latitude in the case of iSHAI version 2016 netCDF

Other reference:

This is the recommend way by other authors and tools. As an example, it is shown in the Figure 2 the definition of one variable in the presentation of Evan Manning, NASA JPL on the NASA Sounder Science Team Meeting (September, 2016). The ***ancillary_variable*** attribute has a different meaning than the georeferencing as can be seen in the Figure 2. It can be seen that the georeferencing is made using the way as in the McIDAS-V case using the attribute ***coordinates***.

Variable "antenna_temp"

```
float antenna_temp(atrack=135, xtrack=96, channel=22);
  :units = "Kelvin";
  :ancillary_variables = "antenna_temp_err";
  :valid_range = 0.0f, 400.0f; // float
  :long_name = "antenna temperature";
  :standard_name = "brightness_temperature";
  :coordinates = "lon lat";
  :description = "Calibrated scene brightness temperature for each ATMS channel and beam position. This
  :_FillValue = 9.96921E36f; // float
  :coverage_content_type = "physicalMeasurement";
```

Figure 2: Example of variable definition in the presentation of Evan Manning, NASA JPL on the NASA Sounder Science Team Meeting (September, 2016)

Then the issue on how to get a good georeferencing of variables (in the sense of correct link the (latitude, longitude) matrices to the variable data not in the sense to get a precise navigation) should be revised and tested with other users, tools and future tools developers. In case of need, this attributes could be added later.

2.1.1 YAW MANEUVERS

In the ATBD of IRS L1 processing is commented that yaw maneuvers will be made every six months. From the point of view of image processing, it would be desirable that the order of the data matrices in the L1 and L2 files be always equal; adding the adequate rotation in some step of the IRS L0=>L1 processing in the inter equinox period that needs it.

If this operation is performed in the EUMESAT ground segment before dissemination or the download from UMARF, it would save many headaches in the users. Of course, a flag indicating that has been rotated the data matrices in the corresponding inter equinox period could be added to the files.

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	**Code:** NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 **Issue:** 1.0 **Date:** 8 November 2017 **File:** Technical_note_comments_IRS_netcdf_structure.docx **Page:** 14/28

In case the rotation is not made, it should be needed before to approve the IRS L1 netCDF format to make tests with as much tools as possible and check that they could accept 4D arrays with latitude and longitude with the matrices rotated and not rotated.

2.1.2 Latitude and longitude arrays inside on netCDF-4 groups.

The longitude and latitude arrays are necessary since in the case of IRS L1 and L2 the projection to a GEO grid will not be performed.

Another issue related to the longitude and latitude fields is the problem that they are related to the navigation in the LWIR band. It has been commented that there is a shifting between the navigations of both bands and several alternatives have been discussed when solving it. Among the alternatives were: provide a global displacement, provide displacement errors, etc.

In the proposed structure in the sample 2016 IRS L1 netCDF files, the longitude and latitude fields are repeated in the two bands LWIR and MWIR, so one solution is to store values of the differenced in the longitude and latitude in the two bands. There would be no problem later in grouping the two bands using the same (column, element) to access the global spectrum.

In the case it will be stored a single longitude and latitude array for the two bands, the solution would be to place these fields in the root group; since I have the doubt of whether the variables of a group can be associated as dimension when the fields that support the dimensions are in a different group. See chapter 2.8 on read of netCDF group's issues.

2.1.3 Proj.4

Almost all Python-based tools use or include a module called proj.4 (among them are CDO, GDAL, etc).

This proj.4 module that can be used from Python for example to reproject an array that has longitude and latitude associated arrays to a series of predefined standard projections. Thus, the correct implementation and association of IRS fields with the navigation arrays (longitude and latitude) is of great importance.

2.1.4 Inverse of (longitude, latitude) to (row, column) on IRS dwell array.

For specialized reprojection operations (as in NWP or in the NWC SAF software) it should be provided a function that using the set the navigations and orbital parameters in the dwell could provide direct and inverse conversion between row, column and longitudes and latitudes.

2.2 SCALE AND OFFSET ISSUE

In sample 2016 IRS L1 netCDF files, *radiance_offset* and radiance_scale_factor are declared as independent variables and they are declared as arrays. This does not follow the standard way.

```
/data/lwir/measured/radiance_offset           H5T_FLOAT [817]
/data/lwir/measured/radiance_scale_factor     H5T_FLOAT [817]
```

In the standard way, the *scale_factor* and *add_offset* are attributes of every variable and the name of the attribute is always the same and they does not include the name of the field.

In the example shown in Table 5, *scale_factor* and *add_offset* are attributes and they are used to store the field TPW of the NWC iSHAI product `ishai_tpw` as an integer 2D array (2 bytes per pixel) instead of a 2D float array (4 bytes per pixel). The TPW values are stored as centesimal of mm of the precipitable water and they are first multiplied by 100 before to store in the netCDF; the *scale_factor* and *add_offset* are used to reconstruct the values after reading them using the standard procedure:

Data_values = *scale_factor* * data_field + *add_offset* *Eq. (1)*

```
ushort ishai_tpw(ny, nx) ;
    standard_name = "ishai_tpw" ;
    long_name = "iSHAI Total Precipitable Water" ;
    scale_factor = 0.01f ;
    add_offset = 0.f ;
    FillValue = -1US ;
..
```

Table 5: Example of definition of one variable with the attributes scale_factor *and* add_offset.

Related to this issue there is an additional comment, in the sample IRS L1 netCDF files, *radiance_offset* is an array fixed to a *missing_data* value (9.96921E36) and I think *radiance_offset* should be fixed to 0.

If what is intended is to skip applying the Equation 1 for the conversion of the archived data to the true value when the *scale_factor* and *add_offset* are not needed, I think it would be better simply not write the corresponding attributes.

When I tried to convert netCDF-4 to netCDF-3 using the CDO application, it did not work due to the bad specification (using of this kind of missing_values) of the scale and offset attributes; as example this happens with netCDF L2 files from the NRT IRS proxy experiment.

In the classical procedure to manage netCDF files:

- First, are read the data and the scale_factor and add_offset attributes
- Then, it is made a checking process to identify which attributes are not present or out of range
- taking into account the present attributes the real data values are obtained through the options:
    1. copy the data if the scale_factor and add_offset attributes are not present or out of range
    2. multiply only by the scale
    3. add just the offset
    4. apply the general formula data = scale * field + offset if the scale_factor and add_offset attributes are present

On other hand, in a Python program provided by Siebren de Haan (KNMI) the read of same netCDF netCDF L2 files from the NRT IRS proxy experiment with the *scale_factor* and *add_offset* attributes fixed to some missing data values was done without any problem in just a line with a Python library. I suppose that internal Python functions make the process above and allow to read internally with a single line. At the same time, this implies that not to follow a standard structure in the IRS netCDF files could compromise the fast read and management of IRS netCDF files by languages as Python or standard libraries of user's tools.

In the sample IRS L1 netCDF file the variables have not the attribute **_FillValue**. The use of this attribute is important to mask out pixels not processed (as example cloudy pixels on IRS L2 products) or with very much degraded quality. See Table 5 (ncdump de iSHAI example)

Other related attribute is **valid_range** the use of this attribute could users to advice users of out of range values.

2.3 HIGH SPATIAL RESOLUTION BROAD BAND IMAGES (480X480 IMAGES AT 1 KM PER BAND) FOR THE DWELL

In the October 2017 IRS MAG workshop, the inclusion of the 480 by 480 pixel images resulting from integrating in the whole band was approved for navigation and control purposes. The issue is how to associate the navigation of a 480 by 480 matrix with 160 x 160 pixel longitude and latitude matrices and to avoid the need to include additional matrices with the longitude and latitude of 480 by 480 pixels.

There is also the issue on how to write the necessary parameters to describe the modified Planck function that is needed to perform the conversion of brightness temperatures to radiances and vice versa. In the case of IRS spectra there is not needed to use a modify Planck function but for these subsample images obtained by the integration on the whole IRS bands it is needed. This should be done using the same way that the necessary parameters to describe the modified Planck function are write in the MTG-FCI netCDF files. It is one of the pending tasks of the satellite community. If a convention were created, it would make it easier for tool developers to include modified Planck routines in a greater number of tools and programming languages such as Python. Thus the conversion from radiances to brightness temperature could be included in more geophysical routines and tools.

2.4 WAVENUMBER DIMENSION ON IRS L1 NETCDF FILES.

Another issue that I disagree with the format of the structure proposed in the sample 2016 IRS L1 netCDF files is the fact that it provides only three parameters (`start_wavenumber`, `end_wavenumber`, `wavenumber_step`) to describe the wavenumbers in the band. See Figure 3.

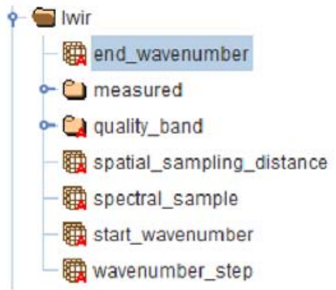


Figure 3: Parameters for wavenumber description in in the sample IRS L1 netCDF files

Following this specification users must reconstruct the array with the wave numbers in memory after the read of the data. In my opinion this is an error and it is better to add also the complete array with the wavenumbers and declare this wavenumber array as other additional dimension.

This way is made in the McIDAS-V test netCDF file; and this is the structure that has been used for the generation of images in the Figure 1.

In this way, the array with the spectra in an IRS L1 band would be an array with 4 dimensions (`wavenumbers,longitude,latitude,time`). This would facilitate the operations on user side of EUMETCast and in addition it would allow later that standard tools will realize the tasks of extraction, combination etc. of the arrays with the spectra.

To design this structure is not enough. It must also needed to contact the developers of tools in order they include the management of this wavenumber dimension in the future in their tools. As example, currently CDO supports as a third dimension z but does not provide support when the vertical dimension is in pressure levels (or I have not be able to make it). In the same way, it should be done for the wavenumbers, pressure levels or hybrid levels and the Principal Components cases.

In this way, when the user will require to perform a simple extraction of a set of levels, a set of wave numbers or a series of Principal Components they could use the same tool; that will also allow at the same time to perform tasks such as internal compression, change of version, combination of several slots in the only file, reprojection, etc.

2.5 IRS L1 AND L2 FILE NAME CONVENTION

The string with the name of the sample 2016 IRS L1 netCDF files are very large but does not contain useful information for the fast and early navigation of the dwell.

IRS L1b filename:

W_XX-EUMETSAT-Darmstadt,SND+SAT,MTS1+IRS-1B-SSS-x-ZI-GEO-CHK-BODY--DIS-NC4E-C-EUMT_20160122140000_IDPFS_DEV_20130305000000_20130305000000_N__T_0000.nc

PC filename:

W_XX-EUMETSAT-Darmstadt,SND+SAT,MTS1+IRS-1B-PC-x-ZI-GEO-CHK-BODY--DIS-NC4E-C-EUMT_20160122140000_IDPFS_DEV_20130305000000_20130305000000_N__T_0000.nc

In my opinion filenames should just provide:

1. the LAC region
2. the number of the dwell with the approximate position to the nearest dwell in the reference map of dwells based on map similar to the shown in Figure 4.
3. the nominal time
4. the start time and end observation time
5. the extension '.nc'

Thus, my proposal for IRS L1 filenames should be similar to:

W_XX-EUMETSAT-Darmstadt,MTG-IRS1-PC-LAC4-dwell036_20160122T140000-_20160122T140521-to-20160122T140541.nc
W_XX-EUMETSAT-Darmstadt,MTG-IRS1-L1B-LAC4-dwell036_20160122T140000-_20160122T140521-to-20160122T140541.nc

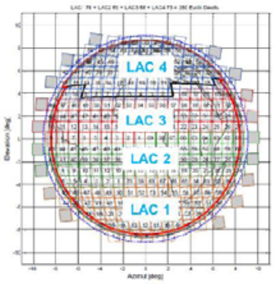


Figure 4: Example of map of approximated positions of LAC regions and dwells.

2.6 INTERNAL VS EXTERNAL COMPRESSION ISSUE.

As I commented some times in previous MIST and MTG-IRS Advisory Groups presentations, it should be discussed the issue of internal compression or external compression of netCDF files.

The use or not of internal compression is a basic decision on the design of netCDF and also it will have a lot of implications not only in the operations but also in science. We must think that our decisions today will have a lot of implications in future in the user's operations; when in future we will design some developments over large regions or wide period is not the same the need to transfer from the archive and made external decompression followed by open&read files that just transfer from archive and open&read files.

The variables in the sample 2016 IRS L1 netCDF files are not internal compressed. To activate internal compression it is easy in the case of Fortran and C when the netCDF file is created for the whole netCDF file.

Note: In the case of IDL-8.4 it is possible to activate the internal compression when every variable is defined (this way you can activate internal compression for one and no compression for other).

I made some variants of the IDL program to write in one netCDF-4 file just the LWIR radiance, wavenumber, longitude and latitude variables and to test the compression rate using internal compression versus writing uncompressed files and later execute the gzip compression tool with same compression level. I have made test writing LWIR radiance as float (4 bytes) or as short integer (2 bytes).

As can be seen in Table 6 the sizes are similar or better in the case of internal compression with the great advantage that the files can be used just after reception. In the case of external compression, it is needed before to use the file decompressed it. **To avoid the external compression/decompression will be a great benefit for users and EUMETSAT.**

Size (bytes)	Size of uncompressed files	Comparison of the size with the uncompressed case (%)
83877381	IASI_mcIDAS_rad_lwir_float_cxy_20130305000000_20130305000000_N__T_0000.nc	100
42046947	IASI_mcIDAS_rad_lwir_int_cxy_20130305000000_20130305000000_N__T_0000.nc	100
	Size of files with internal compression	Comparison of the size with the uncompressed case (%)
43777674	IASI_mcIDAS_rad_lwir_float_cxy_gzip_20130305000000_20130305000000_N__T_0000.nc	52
30928849	IASI_mcIDAS_rad_lwir_int_cxy_gzip_20130305000000_20130305000000_N__T_0000.nc	74
	Size of files with external compression => gzip.exe of file	Comparison of the size with the uncompressed case (%)
45701220	IASI_mcIDAS_rad_lwir_float_cxy_20130305000000_20130305000000_N__T_0000.nc.gz	54
34237628	IASI_mcIDAS_rad_lwir_int_cxy_20130305000000_20130305000000_N__T_0000.nc.gz	81

Table 6: Comparison of the size of uncompressed, internal compressed and external compressed files for just with the LWIR radiance, longitude, latitude and wavenumber variables in the case radiance(nchanel, nx, ny).

Related with this, it is the issue to store IRS L1 spectra in a band as ***radiance***(`nx,ny,nchannels`) or ***radiance***(`nchannels,nx,ny`) as it can be seen in table 7 with the sample IRS-L1 file it is better (`nx,ny,nchannels`) but the compression rates could change with the addition of data noise. In the Figure 1 the fields have not added noise in the data. More tests should be made with this issue.

Size (bytes)	Size of uncompressed files	Comparison of the size with the uncompressed case (%)
83877365	IASI_mcIDAS_rad_lwir_float_xyc_20130305000000_20130305000000_N__T_0000.nc	100
42046931	IASI_mcIDAS_rad_lwir_int_xyc_20130305000000_20130305000000_N__T_0000.nc	100
	Size of files with internal compression	Comparison of the size with the uncompressed case (%)
29305903	IASI_mcIDAS_rad_lwir_float_xyc_gzip_20130305000000_20130305000000_N__T_0000.nc	35
23122409	IASI_mcIDAS_rad_lwir_int_xyc_gzip_20130305000000_20130305000000_N__T_0000.nc	55

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	Code: NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 Issue: 1.0 Date: 8 November 2017 File: Technical_note_comments_IRS_netcdf_structure.docx Page: 19/28

	Size of files with external compression => gzip.exe of file	Comparison of the size with the uncompressed case (%)
33485387	IASI_mcIDAS_rad_lwir_float_xyc_20130305000000_20130305000000_N__T_0000.nc.gz	40
26245314	IASI_mcIDAS_rad_lwir_int_xyc_20130305000000_20130305000000_N__T_0000.nc.gz	62

Table 7: Comparison of the size of uncompressed, internal compressed and external compressed files for just with the LWIR radiance, longitude, latitude and wavenumber variables in the case (nx, ny, nchanel).

I have made also some tests using the netCDF L2 files from the NRT IRS proxy experiment to write arrays of FG and IASI L2 for all pixels in order (`60,nlines,n_hybrid_levels`) and (`n_hybrid_levels,60,nlines`) and compare the compression rate. Since the files were nearly empty (few processed pixels) the results were not significant.

The tests should be repeated with noise added for IRS L1 and make tests (`160,160,n_channels`) versus (`n_channels,160,160`).

Same for PC IRS L1 and make tests (`160,160,n_PC`) versus (`n_PC,160,160`). Same for IRS-L2 files.

Note: In the preparation of the converter, it was used the ECMWF IDL (version 8.4) to write IRS files with internal compression for making some tests to investigate the internal compression. Now I cannot to repeat because the IDL is not possible to use in the ecgate machine at the ECMWF and I have must to use GDL. GDL is not as powerful as IDL and then it will better to make the tests with Python in future.

Other issue related with the internal compression is the algorithm used to make it. I think that the best option is to use internal compression of the disseminated IRS-L1 and IRS-L2 netCDF files with one standard gzip internal compression algorithm. The gzip internal compression is implemented in all the languages that support netCDF library and then any tool that can read netCDF-4 should be able to open&read the IRS L1 and L2 netCDF-4 files. The price to use not standard internal compression as the proposed for the MTG-FCI L1 files is that likely the functions and libraries will be available for only few languages and tools.

As example, the use of a nonstandard (most of all proprietary) internal compression for the MTG-FCI L1 files will be headache source for national meteorological users in the near future till the JPEG-LS (CharLS) internal compression is declared standard and sources and tools start to incorporate it in ordinary tools (if some day occurs!); there is also the issue of the availability of public JPEG-LS (CharLS) internal compression libraries for Python or Java languages.

2.6.1 Principal Components (PC) compression issue.

As can be seen in Figure 5, in the case of Principal Components (PC) sample file, it has been proposed to store as array of 1 or 2 o 3 o 4 bytes by pixel.

I think that this is an error and as can be seen in Figure 6 taking in account that PC images are nearly empty, the compression rate as image should be great and it should be studied if the internal compression in one common array PC(`n_pc,nx,ny`) or PC(`nx,ny,n_pc`) has a similar size to the size of the arrays proposed in the Principal Components (PC) sample file.

NWC SAF AEMet	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	Code: NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 Issue: 1.0 Date: 8 November 2017 File: Technical_note_comments_IRS_netcdf_structure.docx Page: 20/28

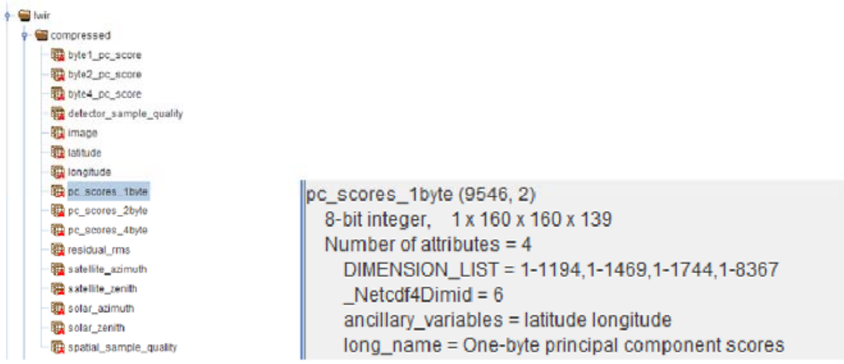


Figure 5: NetCDF structure of PC filename showing pc_scores as 1 or 2 or 3 or 4 bytes arrays.

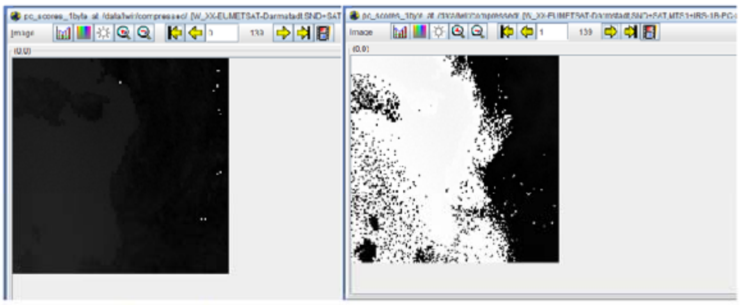


Figure 6: Example of 1st and 2nd PCs of 1 byte1_pc_score in the PC filename showing that images are near empty and the histograms shown that near all pixels have the same value => very high internal compression rate.

2.7 NOT TO USE NEAREST NEIGHBOUR INTERPOLATION OF NWP FILES FOR IRS L1 AND L2 PROCESSING AND DEFINE A SET OF DENSE FIXED PRESSURE LEVELS FOR IRS-L2 PROFILES

Being in hybrids levels ECMWF the profiles the netCDF L2 files from the NRT IRS proxy experiment and in the IRS L1 ATBD document, it can be deduced that EUMETSAT have chosen the nearest neighbour in space and time interpolation to build the generation of a priori NWP profiles for IRS L1 and L2 processing.

As I have already mentioned some times in previous presentations, nearest neighbour interpolation make sense from the point of view of operations (as it facilitates the process) but nearest neighbour interpolation may create false gradients between neighbouring pixels and jumps between consecutive images of IRS L1 and L2 in future. As example, this will have impact on winds calculated using the optical flow algorithm or in the skin temperature estimations.

In the clear air NWC SAF product iSHAI, it is made a triple interpolation for perfect collocation in space and time of NWP and SEVIRI BTs at SEVIRI pixels:

- The process use the GRIB files from the numerical model just before and after the date of the image on hybrid levels or in fixed pressure levels.
- First the nearest previous and posterior GRIB files to the date of the image are searched

- The data from nearest previous and posterior GRIB files are read into memory using GRIB API library,
- Then it is performed for each NWP point the interpolation/extrapolation to a fixed number of pressures (in my case RTTOV pressure levels) for the prior and posterior profiles.
- Then, it is made the time interpolation. It is performed at the date of the image at the RTTOV pressure levels on NWP pixel projection and stored in memory.
- In execution time, it is made spatial bilinear interpolation to each satellite pixel being processed.

Thus a perfect temporal and spatial collocation between IRS pixel and the ECMWF is achieved. In the October 2017 Workshop discussion was recommended to make also the slant correction.

This way, it saves execution time since the spatial interpolation is only made in the processed pixels. This way, it also saves memory since only is needed to allocate memory at the NWP model (in this case the ECMWF) projection. At the time being, ECMWF spatial resolution is near 16 km, which is worse than the IRS resolution and by this fact the size of the arrays to store the NWP model in memory are lower than for the IRS at the full disc IRS projection. As EUMETSAT Secretariat can design a centralized processing algorithm with powerful CPUs it could be designed a pre-processing for the triple interpolation for perfect collocation in space and time of NWP and IRS BTs at IRS pixels.

Other related question is the supply of the ECMWF GRIB files to EUMETSAT. At least for one standard user, the retrieval from ECMWF MARS of GRIB files every hour in hybrid levels for the T, q, u, v and O_3 profiles at high spatial resolution (0.125º x 0.125º) for the full IRS disc in real time is near impossible. It spent more than one hour to get just the GRIB file for one hour with a size of several GBs; in case to try to retrieve from t+6 to t+24 for every hour and every ECMWF run is something impossible to generate and to download. Since the users would like to have access to same NWP used in the IRS L2 and IRS L1 processing (see PGE00 example in NWC SAF web page or in Martinez 2015) for comparison of the input and output of IRS algorithm, it should be studied a way to share the ECMWF the GRIB files used by EUMETSAT with the users. See the example of iSHAI difference fields in *http://www.nwcsaf.org/ishai_description*.

If ECMWF generates the GRIB files for IRS processing faster (outside of the MARS bottleneck) and they would be publicly available (perhaps restricted to national meteorological services) then at least authorized users could avoid the need to retrieve many times the same GRIB files from the MARS. It could be studied also the process to make to get a subset for the region of interest of these GRIB files for IRS processing in the *ecgate* machine of the ECMWF.

Some attributes with the ECMWF GRIB filenames or the forecast time and date of the ECMWF run in the IRS L1 and L2 processing should be included in the IRS L1 and L2 netCDF files.

2.8 READ OF NETCDF GROUP'S ISSUES.

The creation of deep group trees should be avoided.

The groups in netCDF are a little more complicated to read than in HDF-5; although it is also possible to use HDF-5 API to read deep group trees some authors not recommend it.

The version of netCDF IRS L1 and L2 is netCDF4. Since netCDF4 is built on HDF5, the users could use the HDF5 library to read netCDF4 files. Thus, the read of data for a variable can be made using netCDF or HDF5 functions indistinctly but in the case of variables inside groups it is more easy the HDF-5 way than the netCDF way. As can be seen in Table 8, it can be easier to read directly variables when data are inside groups using the HDF-5 library routines.

NWC SAF AEMet Agencia Estatal de Meteorología | Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files | Code: NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0
Issue: 1.0 Date: 8 November 2017
File: Technical_note_comments_IRS_netcdf_structure.docx
Page: 22/28

netCDF-4 way to read /data/lwir/measured/effective_radiance	HDF-5 way to read /data/lwir/measured/effective_radiance
Open file Open group data Open group lwir Open group measured Read variable effective_radiance	Open file Read variable "/data/lwir/measured/effective_radiance"

Table 8: Comparison of read a variable using the netCDF-4 library and the HDF-5 library.

The reason is that in HDF-5 the functions can read directly the `path+variable`; but in the case of netCDF-4 functions it is needed first open the every group in the `path` and later to read the `variable`. At least with IDL for the case of test IRS L1 the read of groups and variables is direct using HDF-5 `path+name` but using the netCDF API in IDL it is needed to open successively all the groups in the path and at the end read the variable. In other languages could be implemented internally but it is made also.

NWC SAF AEMet	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	**Code:** NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 **Issue:** 1.0 **Date:** 8 November 2017 **File:** Technical_note_comments_IRS_netcdf_structure.docx **Page:** 23/28

3. SUMMARY AND RECOMMENDATIONS

The summary of the main proposed changes in the netCDF structure for sample 2016 IRS L1 netCDF files and later for IRS L2 is the following:

It is recommended to have the **coordinates** attribute in every variable and it should be associated to the latitude and longitude fields as dimensions.

In the case it will be stored a single longitude and latitude array for the two bands, the solution would be to place these fields in the root group.

Remove the *latitude* and *longitude* from the wrong attribute *ancillary_variables*

Remove the *radiance_offset* and radiance_scale_factor independent variables and follow the standard way. In the standard way, the *scale_factor* and *add_offset* are attributes of every variable and the name of the attribute is always the same and they does not include the name of the field.

The *scale_factor* and *add_offset* are values and not arrays.

Add also the attributes **_FillValue** and **valid_range** for every variable.

Use a filename convention for an easy and fast approximated location of the IRS dwells (as example the proposed in chapter 2.5).

The IRS netCDF files should use the internal compression based on actual standards (gzip compression algorithm). Since the internal compression depends on how 3D arrays are stored it is needed also to make tests for the issue to store IRS L1 spectra in a band as **`radiance`**`(nx,ny,nchannels)` or **`radiance`**`(nchannels,nx,ny)`.

Same tests in case of IRS L2 files with the number of pressure levels and in the case of Principal Components (PC) IRS files with the number of PC.

- For 2D array fields the structure should be:

  Data(longitude, latitude, time)

- For 3D array fields the structure should be:

  For IRS L1:

  Data(**wavenumber**, longitude,latitude,time) or Data(longitude,latitude, **wavenumber**,time)

  In the sample IRS L1 file for IRS L1 there is not present the "**wavenumber**" dimension.

  For PC IRS L1:

  Data(**PC_order** ,longitude,latitude,time) or Data(longitude,latitude, **PC_order** ,time)

  For IRS L2:

  Data(**pressure_levels** ,longitude,latitude,time) or Data(longitude,latitude, **pressure_levels** ,time)

  As has been commented in chapter 2.7 instead on hybrid levels it should be used a set of fixed pressure levels for the vertical representation. Thus, for IRS-L2 the vertical dimension should be set to one set of fixed pressure levels. As example, it could be used the 101 levels used on RTTOV or could be proposed a dense fixed pressure set with inclusion of standard pressure.

  See as example the recommendations of Eric Fetzer and Evan Manning in a presentation at the NASA Sounder Science Team Meeting (September, 2016).

Note: some must be do in order to decide Data(pressure_levels, longitude, latitude, time) or Data(longitude, latitude, pressure_levels, time). See internal compression chapter 2.6.

In case of IRS PC L1 files not to use the separation in 1, 2, 3 and 4 bytes of the PCs and test it internal compression using `(n_pc,nx,ny,1)` or `(nx,ny,n_pc,1)` could get similar compression rates that with the separation in 1, 2, 3 and 4 bytes of the PC. See 2.6.1

The use of future standards for internal compression should be avoided because it means that it will needed that tools developers incorporate these standards and it is also needed more time till users could update their tools with these future standards.

The creation of deep group trees should be avoided.

It would be desirable that the order of the data matrices in the L1 and L2 files due to the yaw maneuver be transparent for the users; adding the adequate rotation in some step of the IRS L0=>L1 processing if needed. A flag indicating that has been rotated the data matrices or not due to the yaw maneuver should be also included.

Write the parameters to describe the modified Planck function broad band images (480x480) needed to perform the conversion of brightness temperatures to radiances and vice versa in the same way that for MTG-FCI channels.

To study on how add parameters (as variables or attributes) for tools as Proj.4 or GDAL for easier use if needed in future.

Some attributes with the ECMWF GRIB filenames or the forecast time and date of the ECMWF run in the IRS L1 and L2 processing should be included in the IRS L1 and L2 netCDF files.

3.1 RECOMMENDATIONS

As can be deduced from these experiences, the design of the internal structure of the IRS netCDF files appropriate to the tools or libraries that will be used in near future with IRS it is of vital importance to avoid useless efforts; this also will allow more uninitiated users to easily and simply use IRS data in the future.

For all the reason shown in this technical note, EUMETSAT is called upon to try to include the variables related to IRS (and it would be convenient of the rest of the EUMETSAT instruments and products) in the CF convention. Even better it could be coordinated by CGMS or WMO to get a more general applicability convention for satellite; it can be made following the process described in the Annex A.1. As more general the data more possibilities that tools developers could adequate the developments or upgrade of tools to the update convention since they include a wider market. The users would beneficiate of one wider tool offer.

In my opinion EUMETSAT and later CGMS should create or collaborate in the common netCDF-4 format files design for satellite data and products. It is not enough the CF convention, it is needed to stablish common rules on simple (but transcendental and transversal decision) issues as:

- how to make the georeferencing of satellite data (GEO projection, LEO navigation issues and in this IRS is not reprojected by EUMETSAT to a common grid),
- how to write the date
- How to stablish extra dimensions as pressure levels, wavenumbers, PC (Principal Component) and the order of the dimensions, etc.

Once agree by EUMETSAT the common netCDF-4 format files design it is needed to test it with as many tools and users as possible. Then, it is needed to stablish contact with the tool developers to fix in the EUMETSAT design or that the tools developers incorporate the change needed to manage the EUMETSAT common netCDF-4 format file.

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	Code: NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 Issue: 1.0 Date: 8 November 2017 File: Technical_note_comments_IRS_netcdf_structure.docx Page: 25/28

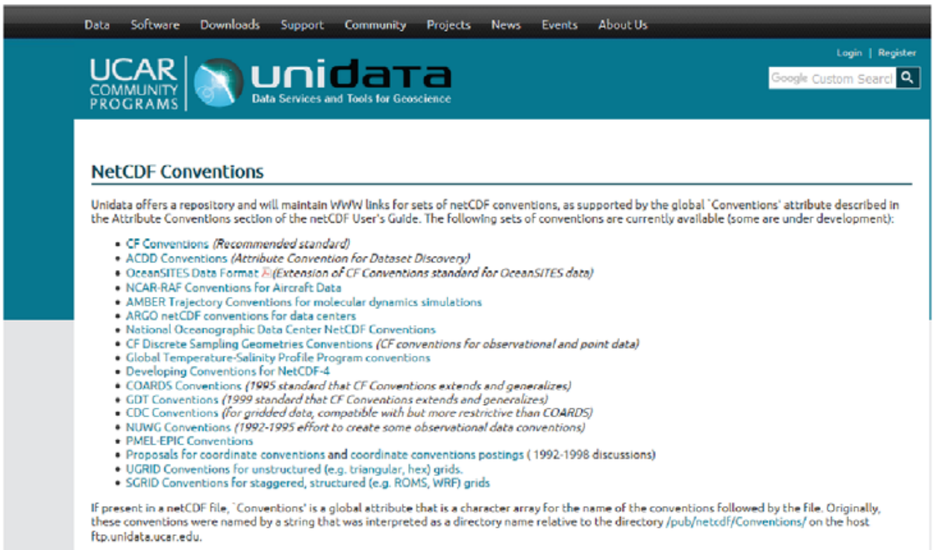
As example, the issue on the use of different byte representation for different PC is a nonstandard feature but some or all tool developers could incorporate in future versions.

It should be studied later the problem of life cycle of the IRS netCDF formats; the need to find ways to make updates of new attributes in future could open new uses of IRS without losing functionality. The major issue should be the modification or redefinition of the dimensions.

4. ANNEX

A.1 Conventions

If present in a netCDF file, `**Conventions**' is a global attribute that is a character array for the name of the conventions followed by the file. The Unidata web page is now the preferred and authoritative location for registering a URI reference to a set of conventions maintained elsewhere.



Other convention web page is http://cfconventions.org/latest.html

A.2.1 Generation of new conventions

It may be convenient for defining institutions and groups to use a hierarchical structure for general conventions and more specialized conventions. For example, if a group named XXX agrees upon a set of conventions for required attributes, attribute names, and netCDF representations for certain discipline-specific data structures, they may describe the agreed-upon conventions in a document associated with the name "XXX", and files that followed these conventions would contain the global attribute

:Conventions = "XXX" ;

Later, if another group agrees upon some additional conventions for a specific subset of XXX data, for example time series data, the description of the additional conventions might be associated with the name "XXX/Time_series", and files that adhered to these additional conventions would use the global attribute

:Conventions = "XXX/Time_series" ;

It is possible for a netCDF file to adhere to more than one set of conventions, even when there is no inheritance relationship among the conventions. In this case, the value of the `Conventions' attribute may be a single text string containing a list of the convention names separated by blank space (recommended) or commas (if a convention name contains blanks), for example

:Conventions = "XXX YYY" ;

NWC SAF AEMet Agencia Estatal de Meteorología	Technical Note: comments on the structure of sample 2016 IRS L1 netCDF files	**Code:** NWC/CDOP2/GEO/AEMET/SCI/RP/IRS_format0 **Issue:** 1.0 **Date:** 8 November 2017 **File:** Technical_note_comments_IRS_netcdf_structure.docx **Page:** 27/28

Typical conventions web sites will include references to documents in some form agreed upon by the community that supports the conventions and examples of netCDF file structures that follow the conventions.

Unidata will maintain this conventions web page and will also make available disk space and web access for additional conventions documents for communities lacking a server on which to publish their own convention hierarchies.

5. REFERENCES

Calbet X.; Martinez M.A., 2016: Future NWC SAF products for the Meteosat Third Generation Infrared Sounder (MTG-IRS). *2016 Convection Working Group Workshop in Florence, Italy.*

Martinez M.A.,; Calbet X, 2016a. Instability indices and water content in clear air NWCSAF product. *2016 Convection Working Group Workshop in Florence, Italy.*

Martinez M.A.,; Calbet X, 2016b. NWCSAF/GEO clear air product version 2016. Evolution to the MTG era. Proc. *Proc. The 2016 EUMETSAT Meteorological Satellite Data User's Conference, Darmstadt, Germany.*

Martinez M.A. 2015. MSG Clear Air Products. Evolution to the MTG era. *See presentation on the NWC SAF 2015 Users' Workshop, February 2015, Madrid (Spain).*

Martinez M.A. 2013a. Improvements in version 2013 of the NWCSAF/MSG PGE13 SEVIRI Physical Retrieval (SPhR) product. *Proc. of the 2013 EUMETSAT Conference, Vienna, Austria.*

Martinez M.A.; et al, 2013b. Innovative ideas for using the hyperespectral level 1 data of the next geostationary MTG-IRS in nowcasting. *Proc. of the 2013 EUMETSAT Conference, Vienna, Austria.*