	<p>Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO</p>	<p>Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG Issue: 2.0 Date: 10 January 2022 File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0 Page: 1/34</p>
---	---	--	--



Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO

NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG, Issue 2.0

10 January 2022

Applicable to

GEO-ASII-NG NWC-088

which is comprised of



GEO-ASII-TF-v2.1

GEO-ASII-GW-v1.1

Prepared by ZAMG

REPORT SIGNATURE TABLE

Function	Name	Signature	Date
Prepared by	A. Jann (ZAMG) for ASII-GW, A. Wirth (ZAMG) for ASII-TF		<i>10 January 2022</i>
Reviewed by	A. Jann (ZAMG) for ASII-TF, A. Wirth (ZAMG) for ASII-GW		<i>10 January 2022</i>
Authorised by	Pilar Rípodas, AEMET SAFNWC Project Manager		<i>10 January 2022</i>

 NWC SAF	 ZAMG	Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG Issue: 2.0 Date: 10 January 2022 File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0 Page: 3/34
---	--	--	--

DOCUMENT CHANGE RECORD

Version	Date	Pages	CHANGE(S)
1.0	21 January 2019	51	First version, for NWC/GEO v2018
2.0	10 January 2022	34	VR for updated versions v2.1 of ASII-TF and v1.1 of ASII-GW (release NWC/GEO 2021)

Table of contents

1. INTRODUCTION	8
1.1 SCOPE AND PURPOSE OF THE DOCUMENT	8
1.2 DEFINITIONS, ACRONYMS AND ABBREVIATIONS	8
1.3 REFERENCES.....	10
1.3.1 <i>Applicable Documents</i>	10
1.3.2 <i>Reference Documents</i>	10
2. VALIDATION OF THE ASII-NG TROPOPAUSE FOLD PRODUCT, V2.1.....	12
2.1 VALIDATION SET-UP	12
2.2 CORRELATION ANALYSIS	14
2.3 THE FRACTIONS SKILL SCORE METHOD	17
2.4 VALIDATION OF THE ASII-TF OUTPUT WITH THE FSS METHOD	18
2.4.1 <i>The FSS for the 25x25-pixel sub-grid</i>	18
2.4.2 <i>The FSS for the 50x50-pixel sub-grid</i>	20
2.4.3 <i>The FSS for the 75x75 pixel sub-grid</i>	23
2.4.4 <i>The FSS for the 100x100-pixel sub-grid size</i>	26
2.4.5 <i>Some remarks on the behavior of the FSS with increasing hit-thresholds</i>	29
2.5 SUMMARY	30
3. VALIDATION OF THE ASII-NG GRAVITY WAVE PRODUCT, V1.1.....	32
4. REFERENCES	34

List of tables

Table 1: List of Applicable Documents.....	10
Table 2: List of Referenced Documents.....	11

List of figures

- Figure 1:** Selected area from the AHI full-disk for the validation of the ASII-TF probabilistic output..... 13
- Figure 2:** IASI L2 derived tropopause height for three consecutive overpasses on 21 March 2021 between 09:20 and 12:50 UTC. The lowest tropopause heights are plotted in red, and the highest-tropopause heights with green dots; height levels in-between in blue. 13
- Figure 3:** IASI L2 derived tropopause gradient for three consecutive overpasses on 21 March 2021 between 09:20 and 12:50 UTC. High tropopause gradient values are plotted with whitish dots. 14
- Figure 4:** Comparison of ASII-TF output (top image; high probability values for a tropopause fold in red) to the IASI L2 tropopause gradients (lower image; high tropopause gradients shown by white dots). AHI data from 21 March 2021 at 12:00 UTC for ASII-TF, and from 09:20 – 12:50 UTC for IASI L2 sounding data..... 15
- Figure 5:** Notation used for the MetOp overpasses over the validation area. Series of numbers in chronological order..... 15
- Figure 6:** Scatterplot of IASI tropopause gradients (x-axis) against ASII-TF probabilities (y-axis) for 25 August 2021 at 12:00 UTC. The red line shows the perfect correlation, the orange ellipse indicates the strong bias towards high tropopause values for low IASI L2 tropopause gradients. 16
- Figure 7:** Line plot of the correlation coefficients for all overpasses (total, blue) and single overpasses (red, grey and orange) corresponding to the first, second and third consecutive overpass. The x-axis shows the validation sample in chronological order. 17
- Figure 8:** 3-D plot of the FSS for a 25x25-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS..... 19
- Figure 9:** Same 3-D plot of the FSS with the 25x25-pixel sub-grid size as in **Figure 8**, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target} line..... 20
- Figure 10:** 3-D plot of the FSS for a 50x50-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS..... 21
- Figure 11:** Same 3-D plot of the FSS with the 50x50-pixel sub-grid size as in **Figure 10**, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target}-line..... 22
- Figure 12:** 3-D plot of the difference between the FSS for a 50x50-pixel and a 25x25-pixel sub-grid. The X-axis and the Y-axis display the ‘hit’ threshold applied to ASII-TF and IASI L2 tropopause height gradient respectively. The vertical axis displays the corresponding FSS difference..... 23
- Figure 13:** 3-D plot of the FSS for a 75x75-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS..... 24

- Figure 14:** Same 3-D plot of the FSS with the 75x75-pixel sub-grid size as in **Figure 13**, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target} line..... 25
- Figure 15:** 3-D plot of the difference between the FSS for a 75x75-pixel and a 50x50-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradient respectively. The vertical axis displays the corresponding FSS difference. 26
- Figure 16:** 3-D plot of the FSS for a 100x100-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS..... 27
- Figure 17:** Same 3-D plot of the FSS for the 100x100-pixel sub-grid as in **Figure 16**, but seen from a different perspective. The red curve marks the ridgeline of the 3-D plot and the black curve the FSS_{target}-line. 28
- Figure 18:** 3-D plot of the difference between the FSS with a 100x100-pixel and a 75x75-pixel sub-grid size. The X-axis and the Y-axis display the hit-thresholds applied to ASII-TF probabilities and to IASI L2 tropopause height gradient respectively. The colors display the corresponding FSS difference. 29
- Figure 19:** 3-D plot of the FSS for a 75x75-pixel sub-grid for 18 March 2021 at 12:00 UTC. The X-axis and the Y-axis display the ‘hit’ threshold applied to ASII-TF and IASI L2 tropopause height gradient respectively. The vertical axis displays the corresponding FSS..... 30
- Figure 20:** Case of 28 December 2021, 1200 UTC. Upper panel: MSG-4 IR10.8 image. Ripples characteristic for gravity waves and comparatively easily spotted in the image are present over Tunisia and the western part of the Gulf of Lion. Lower panel: The probability-of-occurrence of gravity waves, as analysed by the ASII-GW module, v2021, on the basis of the IR10.8 image (the colour table runs from dark blue = 0% to red=100%)..... 32

1. INTRODUCTION

The EUMETSAT’s “Satellite Application Facilities” (SAFs) 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, 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://www.nwc-saf.eumetsat.int>.

1.1 SCOPE AND PURPOSE OF THE DOCUMENT

This document is the Validation Report for the NWC/GEO Automatic Satellite Interpretation – Next Generation Product (PGE17, ASII-NG), NWC/GEO release 2021, comprised of ASII-TF (tropopause fold detection sub-product), v2.1 (chapter 2) and ASII-GW (gravity wave detection sub-product), v1.1 (chapter 3).

This document contains a description of the validation method and the corresponding results for ASII-TF. Concerning ASII-GW, it directs the reader to two recent publications documenting the latest validation efforts.

1.2 DEFINITIONS, ACRONYMS AND ABBREVIATIONS

AHI	Advanced Himawari Imager
AMSU	Advanced Microwave Sounding Unit
ASII	Automatic Satellite Image Interpretation
ASII-GW	Gravity wave detection sub-product of ASII-NG
ASII-NG	Automatic Satellite Image Interpretation – Next Generation
ASII-TF	Tropopause fold detection sub-product of ASII-NG
ATBD	Algorithm Theoretical Baseline Document
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FL	Flight Level
FSS	Fractions Skill Score
GW	Gravity wave(s)
IR	Infrared
IASI L2	Infrared Atmospheric Sounding Instrument Level 2
MetOp	Meteorological Operational Satellite
MSG	Meteosat Second Generation
NWCSAF	SAF to support NoWCasting and Very-Short-Range Forecasting



Scientific and Validation Report
for the “Automatic Satellite Image
Interpretation – Next Generation”
Processors of the NWC/GEO

Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG
Issue: 2.0 **Date:** 10 January 2022
File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0
Page: 9/34

NWP	Numerical Weather Prediction
PIREP	Pilot Report
PGE	Product Generation Element
SAF	Satellite Application Facility
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SW	Software
UTC	Universal Time Coordinated
VIS	Visible
WV	Water Vapour

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 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.

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

Ref	Title	Code	Vers	Date
[AD.1]	Project Plan for the NWCSAF CDOP3 phase	NWC/CDOP3/SAF/AEMET/MGT/P P	1.6	01/12/21
[AD.2]	NWCSAF CDOP3 Project Plan Master Schedule	NWC/CDOP3/SAF/AEMET/MGT/P P/MasterSchedule	1.6	01/12/21
[AD.3]	Configuration Management Plan for the NWC SAF	NWC/CDOP3/SAF/AEMET/MGT/C MP	1.1	15/04/20
[AD.4]	System and Components Requirements Document for the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/SC RD	1.0	01/09/21
[AD.5]	Interface Control Document for Internal and External Interfaces of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/IC D/1	2.0	01/09/21
[AD.6]	Interface Control Document for the NWCLIB of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/IC D/2	2.0	01/09/21
[AD.7]	Data Output Format for the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/DO F	2.0	10/01/22
[AD.8]	Component Design Document for the NWCLIB of the NWC/GEO	NWC/CDOP2/GEO/AEMET/SW/AC DD/NWCLIB	2.0.1	31/07/18
[AD.9]	NWC SAF Product Requirements Document	NWC/CDOP3/GEO/AEMET/MGT/P RD	1.5	01/12/21
[AD.10]	User Manual for the Tools of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SCI/U M/Tools	2.0	10/01/22

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.

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

Ref	Title	Code	Vers	Date
[RD.1]	The Nowcasting SAF Glossary	NWC/CDOP3/SAF/AEMET/MGT/G LO	1.0	20/10/20
[RD.2]	User Manual for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO: Science Part	NWC/CDOP3/GEO/ZAMG/SCI/UM/ ASII-NG	2.0	10/01/22
[RD.3]	Algorithm Theoretical Baseline Document for the “Automatic Satellite Image Interpretation – Next Generation” Processor of the NWC/GEO	NWC/CDOP3/GEO/ZAMG/SCI/AT BD/ASII-NG	1.0	01/09/21
[RD.4]	Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO	NWC/CDOP3/GEO/ZAMG/SCI/VR/ ASII-NG	1.0	21/01/19

Table 2: List of Referenced Documents

		Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG Issue: 2.0 Date: 10 January 2022 File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0 Page: 12/34
---	---	--	---

2. VALIDATION OF THE ASII-NG TROPOPAUSE FOLD PRODUCT, V2.1

The ASII-TF product aims at identifying tropopause folds in the atmosphere with help of satellite and model data. Tropopause folds are strong gradients of the height of the tropopause; they usually coincide with fronts and mark zones where heavy turbulence may occur. ASII-TF provides probabilities of the occurrence of a tropopause fold on a pixel basis.

For the validation campaign described herein, the ASII-TF product has been produced with Himawari-8 satellite data and ECMWF model fields for a pre-defined region of the northern hemisphere (*Figure 1*). NWC/GEO v2021 for the first time offers ASII-TF products tuned to Himawari and GOES-R sensors; a validation of the MSG product was carried out already in [RD.4], hence the focus exclusively on the higher-resolution Himawari product.

In order to have an independent data source for validation, IASI L2 sounding data of MetOp-B and MetOp-C polar orbiting satellites has been analyzed for the selected region and dates. Tropopause height was derived from IASI L2 humidity profiles. Finally, the gradient of the tropopause height was calculated.

To evaluate the agreement of the ASII-TF probabilistic output with the reference dataset, the correlation coefficient between ASII-TF probability value and IASI L2 tropopause height gradient was calculated for all grid points of the validation area where both data sources were available.

In a second approach, the method of the Fractions Skill Score (FSS) was applied. The Fractions Skill Score is a spatial verification metric routinely computed in operational verification suites. It enables the comparison of forecasts of different resolutions against a common ‘truth’ in such a way that high-resolution forecasts are not penalized for representativeness errors that arise from the ‘double penalty’ problem (Mittermaier et al., 2013).

2.1 VALIDATION SET-UP

The data used for the validation of the ASII-TF probabilistic output spans over a period from 1 February to 11 November 2021 comprising 106 single datasets. A dataset consists of the Himawari-8 ASII-TF product output from 12:00 UTC and the matching consecutive MetOp overpasses over the validation area (within an interval of about three hours around noon). The validation dataset has a strong bias towards spring and summer cases.

MetOp-B and MetOp-C overpasses were available twice a day over the validation area, around midnight and around noon. Depending on the satellite flight path, 2 to 3 consecutive overpasses were recorded. For two days, receiving station problems reduced the number of recorded and analyzed IASI L2 data to one overpass and for 33 more days, to two instead of the three evaluated overpasses. Only the noon overpasses were used for the validation.

The evaluation was done for a region of 1570x650 pixels (*Figure 1*) extracted from the Advanced Himawari Imager (AHI) full-disk data and covering Korea, eastern parts of China, Japan and parts of the Pacific Ocean. The region is centered over Japan at latitude 42.06° / longitude 140.7°.

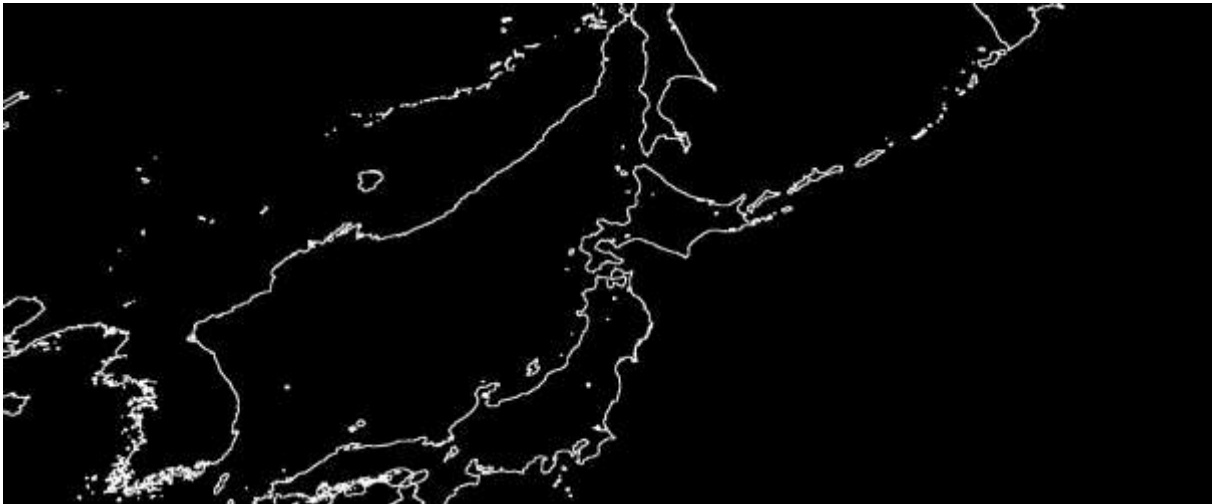


Figure 1: Selected area from the AHF full-disk for the validation of the ASII-TF probabilistic output.

IASI L2 sounding data from MetOp-B and MetOp-C were used as reference dataset for validation. Only mixing-ratio values and attributed height levels were used from the IASI L2 sounding data. The tropopause pressure-level was calculated for each sounding by applying a mixing-ratio threshold of $2.0 \cdot 10^{-5}$ [kg/kg]. As humidity rapidly decreases at the tropopause, this threshold is commonly considered a representative value for locating its height level. In a further assumption, the height of the tropopause is supposed to lie above the 500 hPa pressure level. Missing tropopause levels were interpolated from nearby IASI L2 soundings. Afterwards, the IASI L2 tropopause pressure levels were smoothed with a 5×5 average filter. The gradient of the tropopause height was calculated taking into account the unequal distribution of sounding data points that is due to the instruments scanning geometry (*Figure 2*). The outermost IASI scanlines were discarded after gradient calculation. Plots of IASI L2 derived tropopause heights (e.g., *Figure 2*) and tropopause gradients (e.g., *Figure 3*) were available for visual inspections.

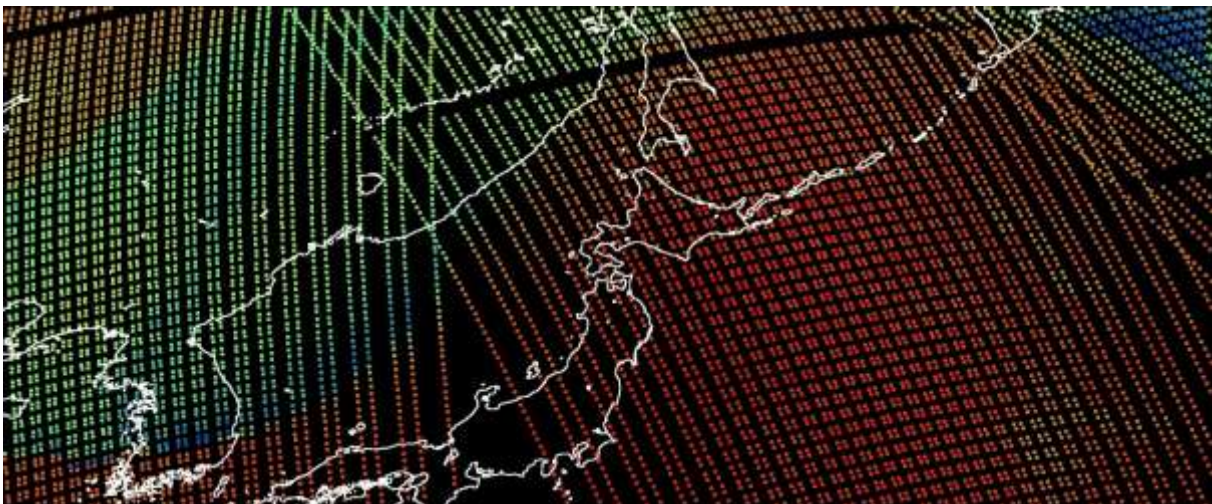


Figure 2: IASI L2 derived tropopause height for three consecutive overpasses on 21 March 2021 between 09:20 and 12:50 UTC. The lowest tropopause heights are plotted in red, and the highest-tropopause heights with green dots; height levels in-between in blue.

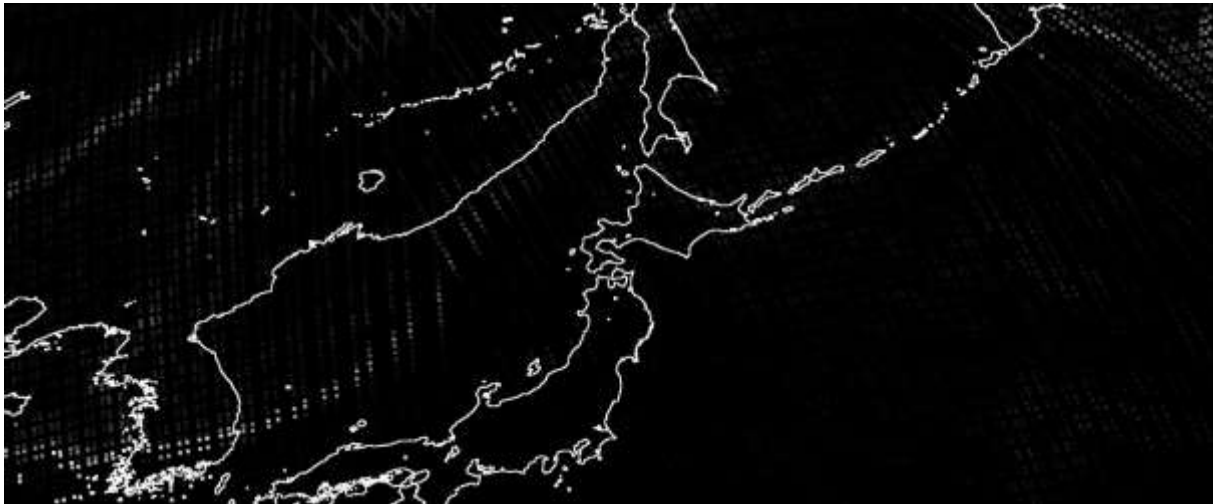


Figure 3: IASI L2 derived tropopause gradient for three consecutive overpasses on 21 March 2021 between 09:20 and 12:50 UTC. High tropopause gradient values are plotted with whitish dots.

2.2 CORRELATION ANALYSIS

The correlation coefficient was calculated for all grid points on the validation area having information on both, the ASII-TF probability value and the IASI L2 tropopause gradient. Correlation analysis was performed for all available overpasses altogether, but also for each single overpass of a given validation sample. For three consecutive overpasses, approximately 8-9000 co-located grid points were available. Accordingly, single overpasses showed less co-located grid points, primarily depending on the geographic position of the satellite swath with respect to the validation area (*Figure 2*).

As can be seen in a direct lineup, IASI L2 soundings have a much coarser horizontal resolution compared to AHI sensor onboard Himawari-8 and hence ASII-TF. Therefore, IASI L2 sounding data only reflect major tropopause folds omitting smaller scale atmospheric disturbances as shown below in a direct comparison of the two outputs (*Figure 4*).

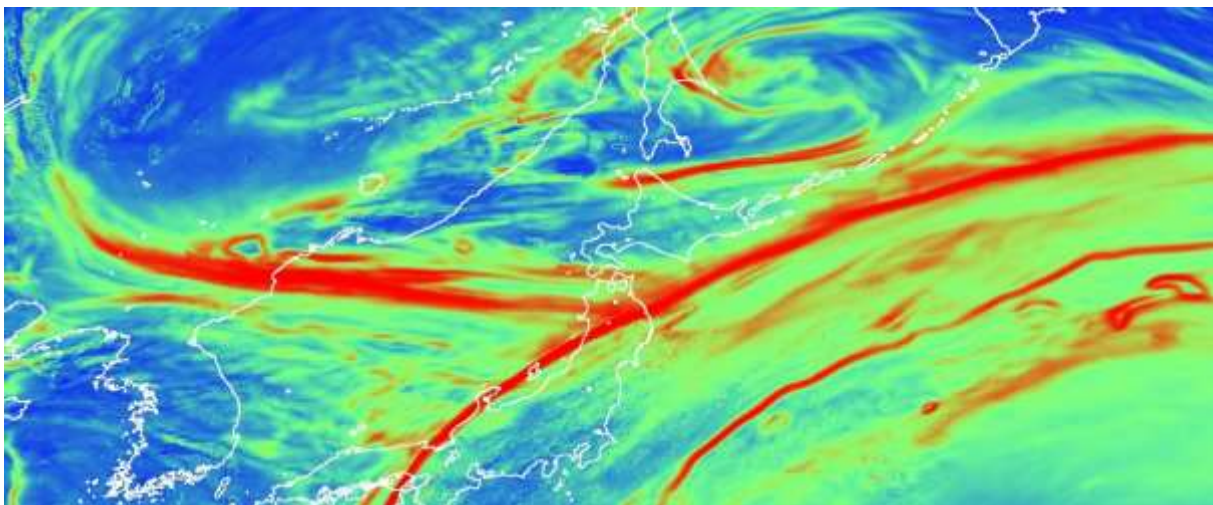




Figure 4: Comparison of ASII-TF output (top image; high probability values for a tropopause fold in red) to the IASI L2 tropopause gradients (lower image; high tropopause gradients shown by white dots). AHI data from 21 March 2021 at 12:00 UTC for ASII-TF, and from 09:20 – 12:50 UTC for IASI L2 sounding data.

Typically, three consecutive MetOp overpasses were recorded over the area of interest close in time to the ASII-TF output (**Figure 5**). From the three overpasses that covered the validation area for a single validation dataset, the overpass #2 showed the best correlation result as it was closest to 12:00 UTC. This is the overpass that was centered over the validation area in most cases.

For a few validation datasets, one (rarely two) of the three single overpasses were not recorded by the ZAMG receiving station, which led to a variable number of summarized single overpasses.



Figure 5: Notation used for the MetOp overpasses over the validation area. Series of numbers in chronological order.

The correlation coefficients have been calculated for all single overpasses (#1, #2 and #3, as indicated in **Figure 5**) over the validation area (94, 105 and 81 single overpasses respectively) as well as for all overpasses together (280 overpasses in total). The shown correlation results are the arithmetic mean of all 106 validation datasets, i.e., for the complete set of overpasses (280) as well as for all ‘first’ (94), ‘second’ (105) and ‘third’ (81) overpasses.

The arithmetic means of the correlation coefficients were:

- For the collection of all ‘first’ IASI overpasses: **0.4035**
- For the collection of all central IASI overpasses: **0.4442**
- For the collection of all ‘third’ overpasses: **0.4001**

The correlation coefficient for all overpasses (i.e., the whole validation dataset) is: **0.424**

The root mean square deviation for overpasses one, two and three are **0.1222**, **0.0864** and **0.1142** respectively. Deviation seems smallest for overpasses that are closest in time to the ASII-TF analysis.

The relatively low correlation coefficients for the ASII-TF output against the IASI L2 tropopause gradients are resulting from the fact that a high-resolution dataset is compared to a low-resolution dataset. The extra information contained in the high-resolution output of ASII-TF is not reflected in the IASI L2 sounding data. This becomes obvious when looking at the scatterplot of *Figure 6*. The plot shows a strong bias towards higher ASII tropopause probabilities when the IASI tropopause gradient is weak (orange ellipse).

A seasonal trend is not visible in the validation dataset as can be seen in *Figure 7*. Outliers in the correlation analysis are often related to a small number of IASI soundings for a specific overpass.

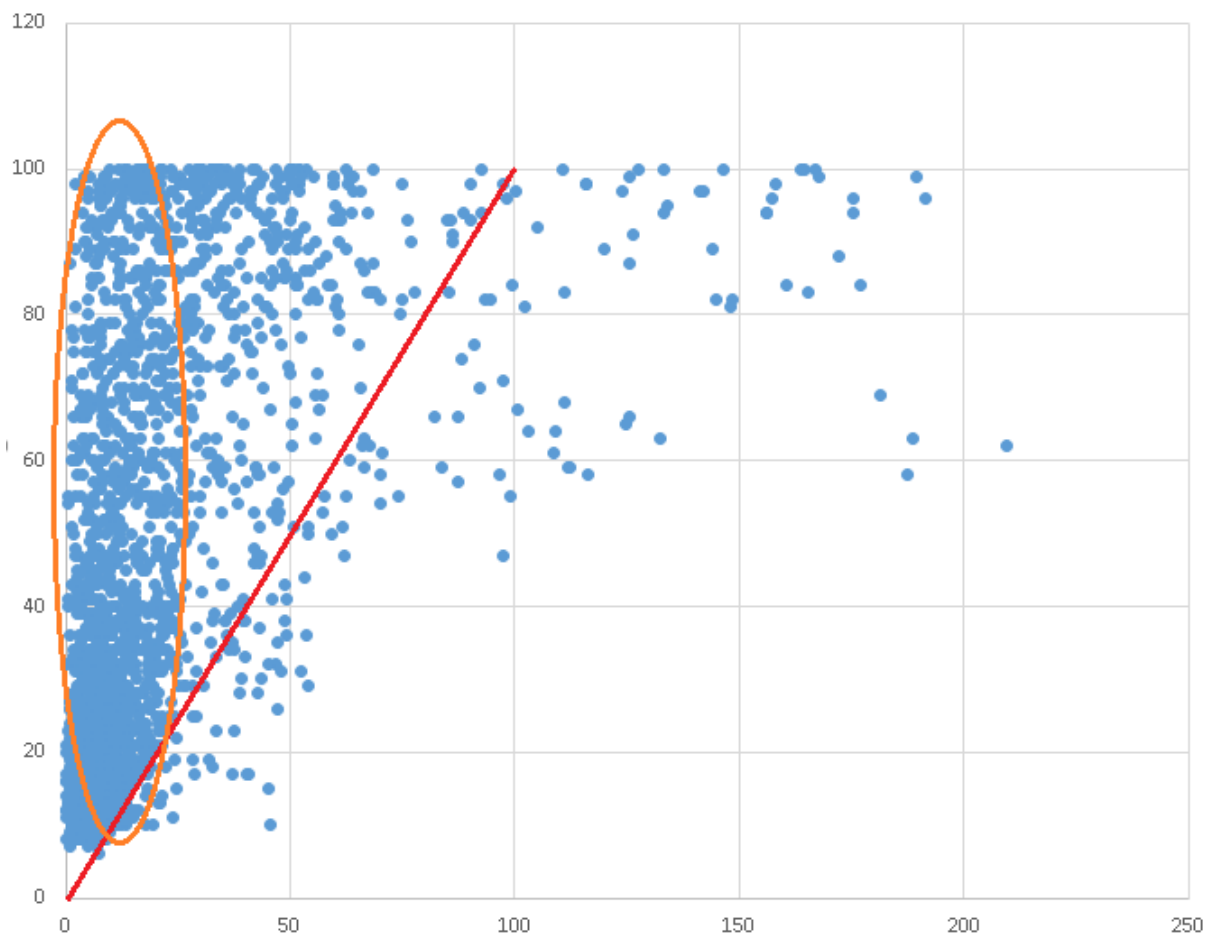


Figure 6: Scatterplot of IASI tropopause gradients (x-axis) against ASII-TF probabilities (y-axis) for 25 August 2021 at 12:00 UTC. The red line shows the perfect correlation, the orange ellipse indicates the strong bias towards high tropopause values for low IASI L2 tropopause gradients.

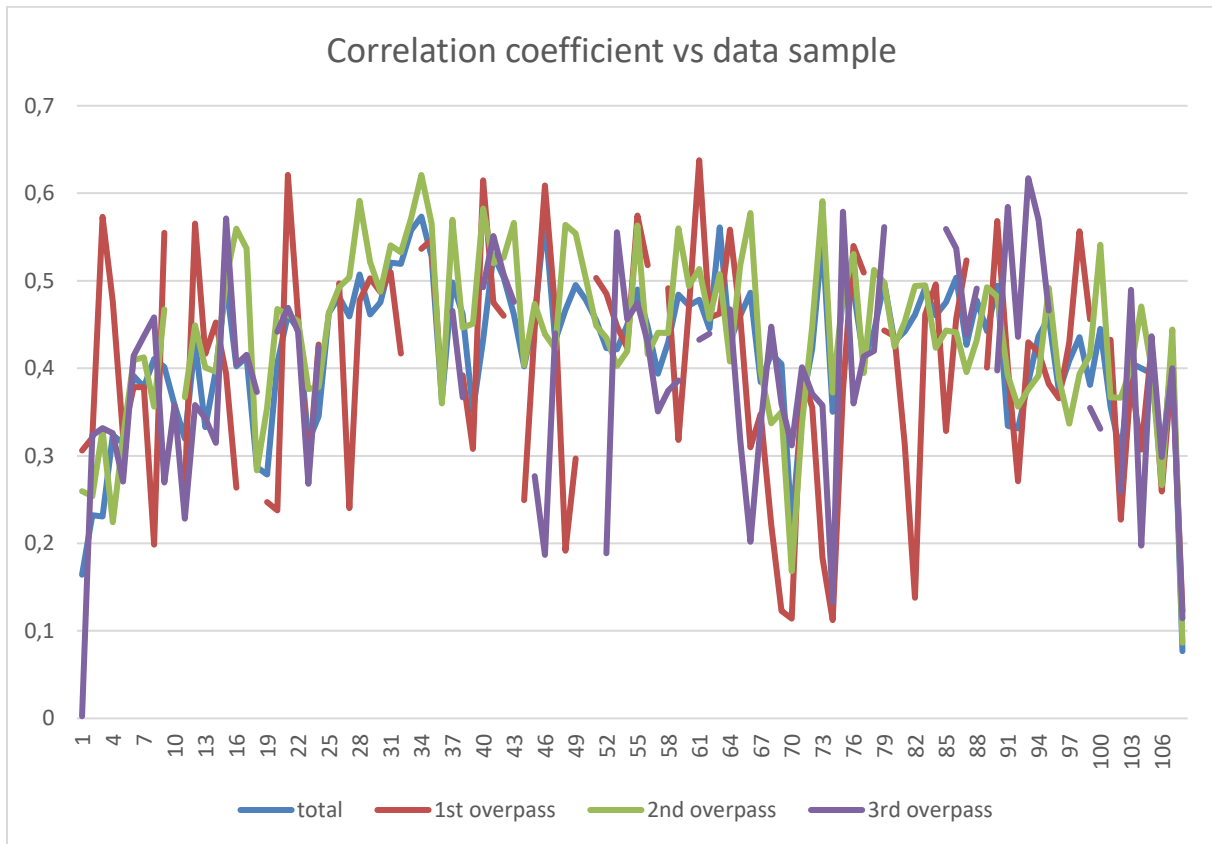


Figure 7: Line plot of the correlation coefficients for all overpasses (total, blue) and single overpasses (red, grey and orange) corresponding to the first, second and third consecutive overpass. The x-axis shows the validation sample in chronological order.

2.3 THE FRACTIONS SKILL SCORE METHOD



The Fractions Skill Score (FSS) method has been introduced by Roberts and Lean in 2008 and has shown good practicability for validating precipitation forecasts that often suffer the double penalty problem of being either predicted at the wrong place or for the wrong point in time. This method is best suited for comparing forecasts and analyses with different horizontal resolution as in our validation example. In order to validate the output of ASII-TF, IASI L2 tropopause height gradient was considered as ‘truth’.

The FSS method consists, simply spoken, in subdividing the validation area into smaller boxes (or fractions) in which the number of ‘hits’ is counted for both the reference data (IASI L2) and the product to be validated (ASII-TF). The location of the hits does not play a role as long as the hit is located inside the same sub-domain. Therefore, two essential parameters have to be defined beforehand: the size of the sub-areas (window size) and the threshold for a hit.

Typically, the FSS is computed for a large range of window sizes and the resulting score plotted as a function of window size. This allows to determine the scale at which the FSS reaches the target skill of

$$FSS_{target} = 0.5 + \frac{f}{2} \quad [Eq. 1]$$

(Faggian et al., 2015), where f is the fractional ‘hit’ coverage over the validation area (hit-area ratio). If f is not very large (within a large domain), a value of 0.5 can be used as a lower limit, and then skillfulness can be assumed for $FSS \geq 0.5$ (Zhao and Zhang, 2018).

		Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG Issue: 2.0 Date: 10 January 2022 File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0 Page: 18/34
---	---	--	---

2.4 VALIDATION OF THE ASII-TF OUTPUT WITH THE FSS METHOD

As mentioned above, two parameters have to be defined beforehand:

1. A threshold that distinguishes a hit from a miss
2. The size of the sub-grid area

Ad. 1.:

While normally, the FSS method is applied to two fields with the same physical quantity, in our validation example, ASII-TF probability values [0-100%] are validated against IASI L2 tropopause height gradient [hPa/km]. To allow the use of a common hit threshold, IASI L2 tropopause height gradients have been rescaled to [0-100] such that value 100 was reached for a tropopause gradient of 200 hPa/km (and everything beyond that threshold was set to 100 as well).

Hit thresholds reaching from 10 to 90 were applied for both data sources, which results in $81^2 = 6561$ Fractions Skill Scores for every test case sample.

Ad. 2.:

Four different sub-grid sizes were used for the evaluation with the FSS method: 25x25, 50x50, 75x75 and 100x100 pixel. The selection of the sub-grid area size implies that small parts of the validation domain were not covered when the boxes are regularly distributed over the validation region (i.e., 1570x650 pixel). The covered areas were:

- 1550x650 for the 25x25 pixel sub-grid (i.e., 1612 fractions)
- 1550x650 for the 50x50 pixel sub-grid (i.e., 403 fractions)
- 1500x600 for the 75x75 pixel sub-grid (i.e., 160 fractions) and
- 1500x600 for the 100x100 pixel sub-grid (i.e., 90 fractions)

The size of the sub-grid is a measure for the horizontal accuracy of the forecast with larger sub-grid areas (fractions) allowing for a stronger displacement between forecast and analysis. This is an essential parameter in our validation as the IASI overpasses rarely fit in time to the available ASII-TF output (i.e., 12:00 UTC).

Note that larger sub-grid window sizes usually lead to better results in the FSS as they allow for larger horizontal shifts between forecast and analysis data. By stepwise augmenting the sub-grid size in this validation example, the intention was to investigate the evolution of the improvement.

The graphs in the following show the average FSS computed over the whole validation dataset (106 cases).

2.4.1 The FSS for the 25x25-pixel sub-grid

The first of the sub-grid areas allows only for a small shift between the ASII-TF output and the IASI L2 tropopause height gradient. The resulting plot of the FSS (*Figures 8 and 9*) can be characterized as follows:

- A low hit threshold gives the highest FSS.
- The ridge line of the 3-D FSS plot corresponds to equal ‘hit’-thresholds for both data sources.
- The FSS drops rapidly when hit-thresholds deviate strongly from each other.
- The FSS drops rapidly for smaller hit-thresholds, to a lesser degree for medium hit-thresholds and again stronger for hit values above 80.

Note that for better visibility in the 3-D plots, a factor of 100 has been applied to the FSS, and that only ‘hit’-thresholds between 10 and 90 were used for FSS calculations.

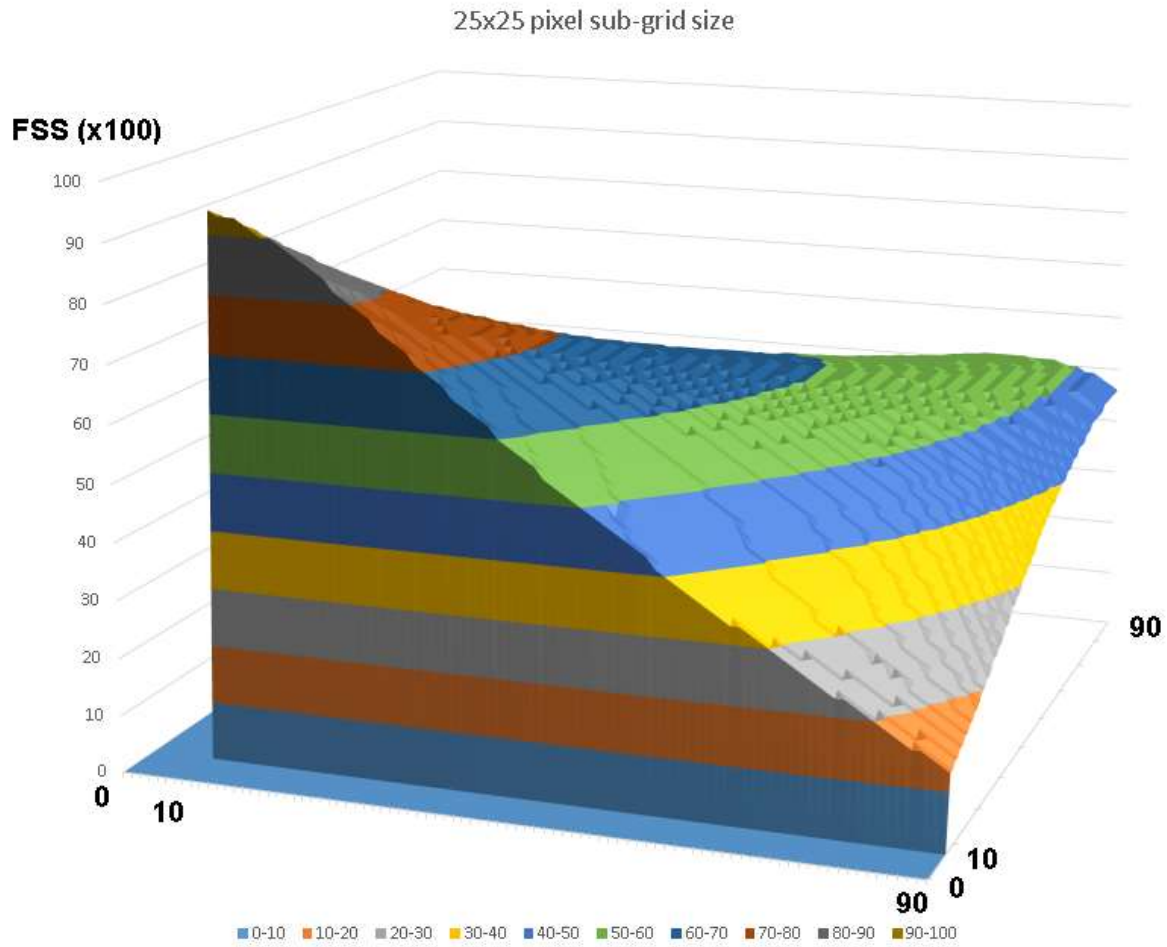


Figure 8: 3-D plot of the FSS for a 25x25-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS.

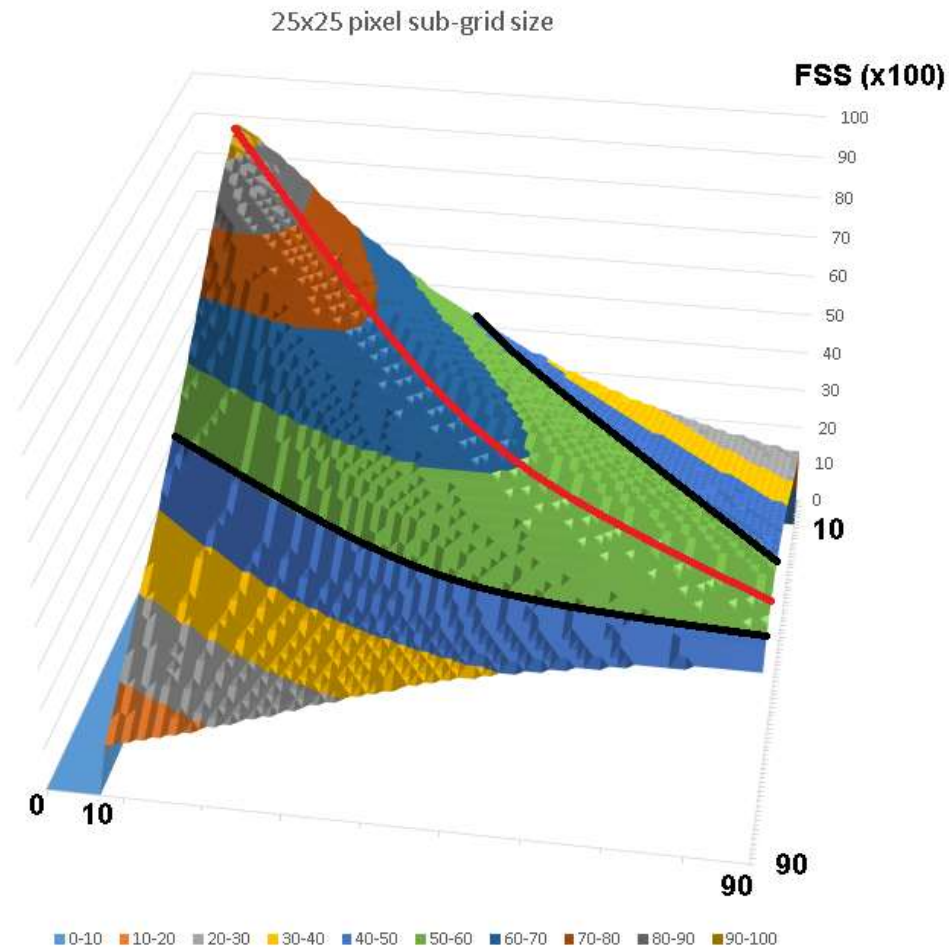


Figure 9: Same 3-D plot of the FSS with the 25x25-pixel sub-grid size as in **Figure 8**, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target} line.

From the changed perspective on FSS in **Figure 9**, it can be seen that there is always a hit-threshold combination that fulfils the FSS_{target} criteria as defined by Mittermaier and Roberts (2010). However, it can also be seen that for higher hit-thresholds, the number of possible threshold combinations for both data sources to fulfil the FSS_{target} criteria is much lower than for smaller hit-thresholds. Note also, that the black line is the FSS_{target} threshold for small hit area ratios f (see **Eq.1**) compared to the validation area. This is certainly not the case for very low hit-thresholds and the line of FSS_{target} should then be well above 0.5 (i.e., 50 in the graph). Considering the high FSS values for low hit-thresholds, this can be taken for granted.

The red line in **Figure 9** spans along the ridge of the 3D plot. It shows the highest FSS values for a given hit-threshold. It is more or less identical to the $x=y$ -axis. However, the red line seems to bend slightly towards lower IASI tropopause height gradients which is caused by the scaling of the IASI L2 tropopause height gradient values.

2.4.2 The FSS for the 50x50-pixel sub-grid

By quadrupling the size of the sub-grid area, one expects higher FSS as the permissible shift between the IASI L2 and ASII-TF tropopause folds can be larger. This is indeed the case when we look at **Figure 10**.

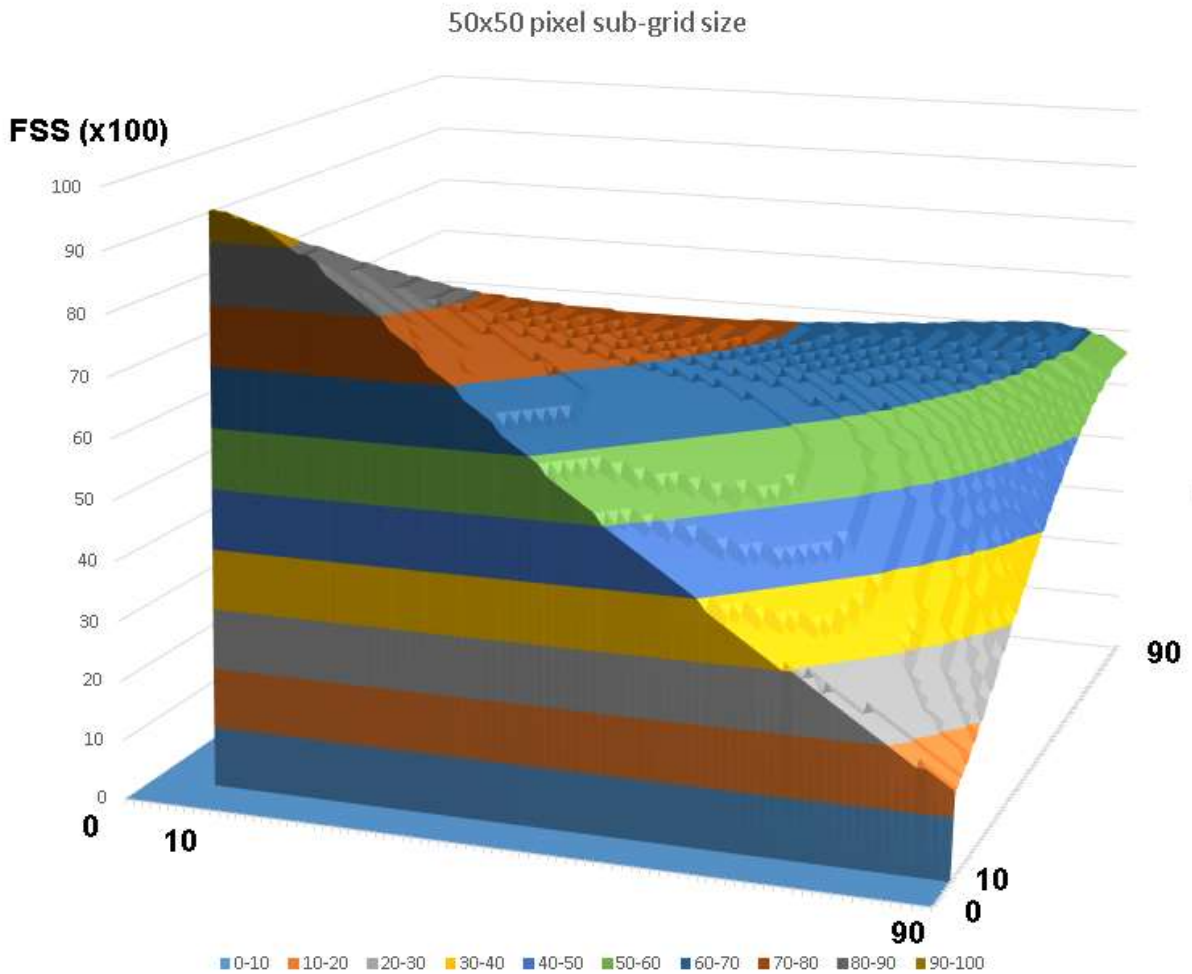


Figure 10: 3-D plot of the FSS for a 50x50-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS.

A change of the perspective on the 3-D plot makes the difference even clearer. Again, the red line marks the highest FSS along the ridge of the 3-D plot. There is always a hit-threshold above the FSS=0.6 line (i.e., 60 in the graph).

The improvement of the FSS is mainly seen for higher hit thresholds, making the slope of the ridgeline less steep compared to the 3-D plot of the smallest sub-grid (see red line in **Figure 11**).

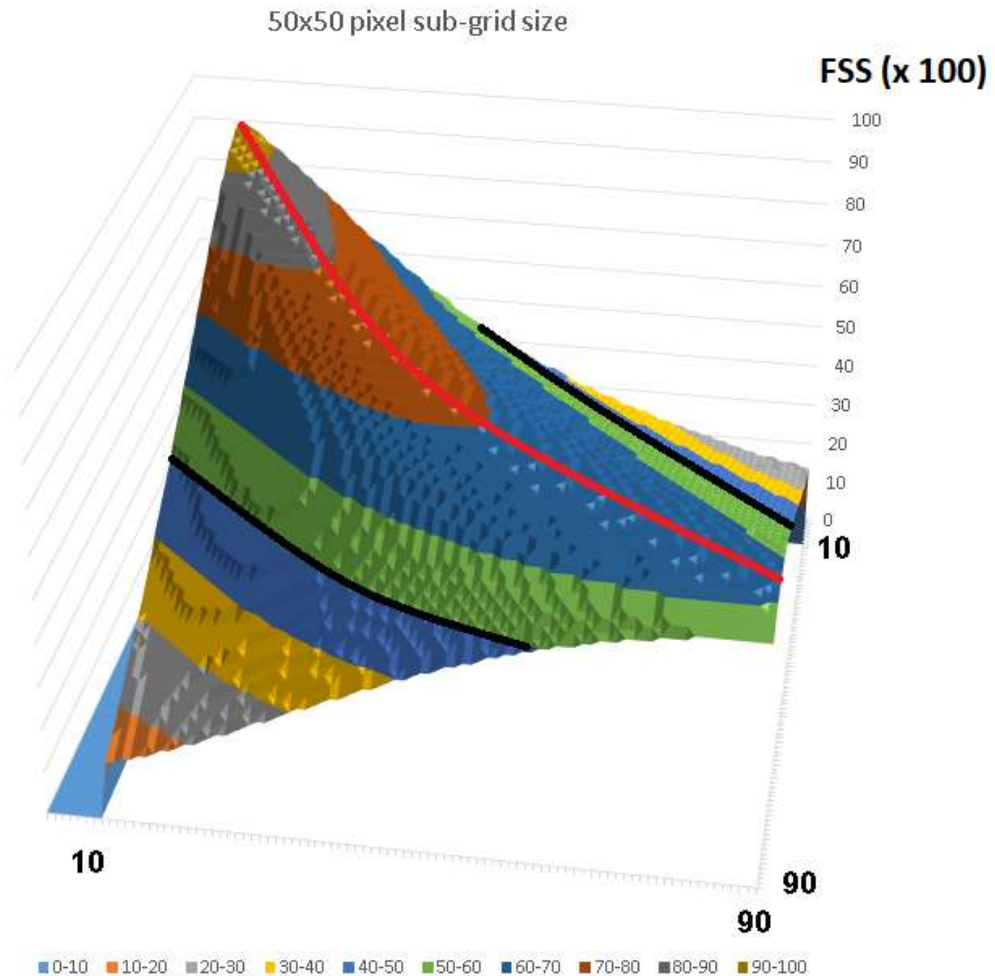


Figure 11: Same 3-D plot of the FSS with the 50x50-pixel sub-grid size as in **Figure 10**, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target} line.

By increasing the sub-grid size, the FSS improves mainly for higher hit-thresholds as depicted in **Figure 12**. This result is expected since a larger sub-grid size allows for stronger displacements between the tropopause folds in ASII-TF and IASI L2. Such horizontal displacements are inherent when comparing data from polar orbiting satellites with data from geostationary satellites due to different scanning times.

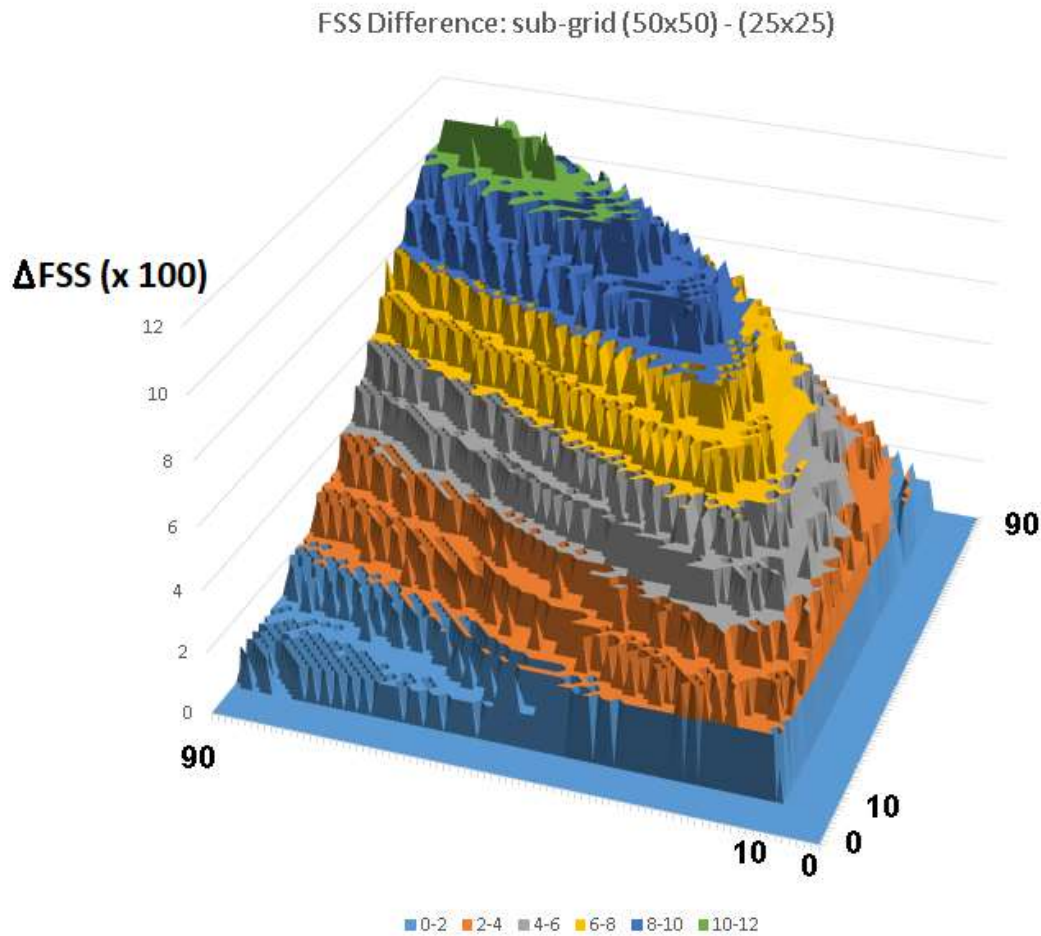


Figure 12: 3-D plot of the difference between the FSS for a 50x50-pixel and a 25x25-pixel sub-grid. The X-axis and the Y-axis display the ‘hit’ threshold applied to ASII-TF and IASI L2 tropopause height gradient respectively. The vertical axis displays the corresponding FSS difference.

2.4.3 The FSS for the 75x75 pixel sub-grid

Augmenting the sub-grid size to 75x75 pixel leads to a further improvement of the FSS (**Figures 13 and 14**). The augmentation is, seen from the area size point of view, less in percentage than the previous steps, but the same from the distance (i.e., shift between forecast and analysis) point of view. Still the improvement of the resulting FSS is considerable, as can be seen in **Figure 15**.

The changed perspective on FSS in **Figure 14** shows that the score is well above 0.5 (= 50 in the 3-D plot) when considering the ridgeline of the plot (red line). This confirms a good agreement of both data sources when using similar hit-thresholds.

As in the previous step from 25x25 to 50x50-pixel sub-grid size, the increase of the FSS is still considerable for higher hit thresholds (**Figure 15**) in the step from 50x50 to 75x75-pixel.

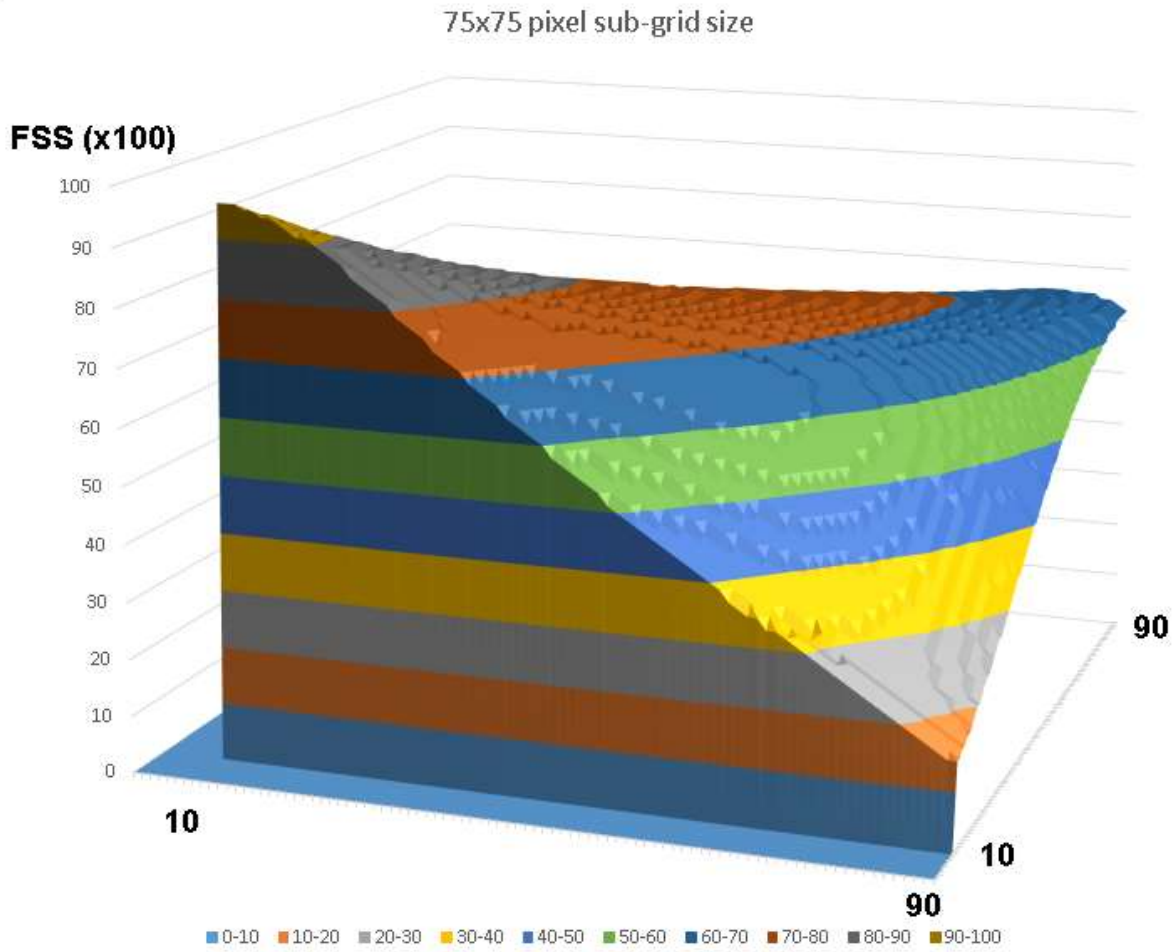


Figure 13: 3-D plot of the FSS for a 75x75-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS.

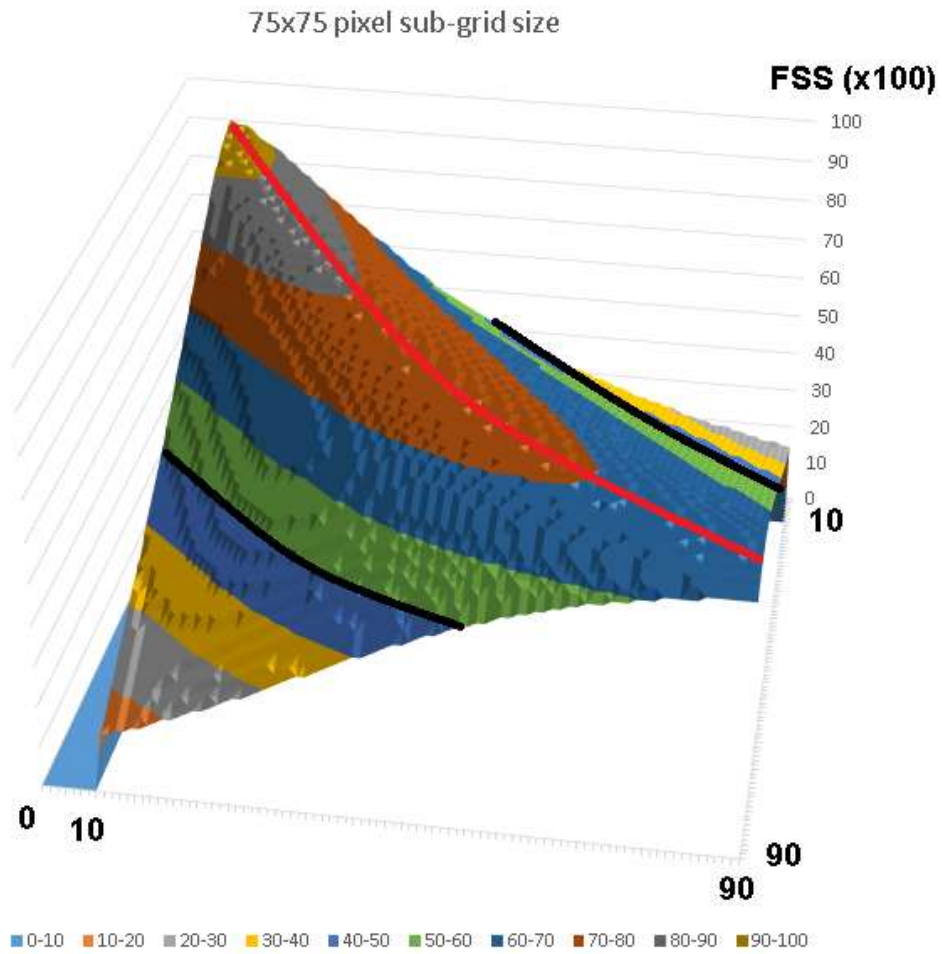


Figure 14: Same 3-D plot of the FSS with the 75x75-pixel sub-grid size as in **Figure 13**, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target} line.

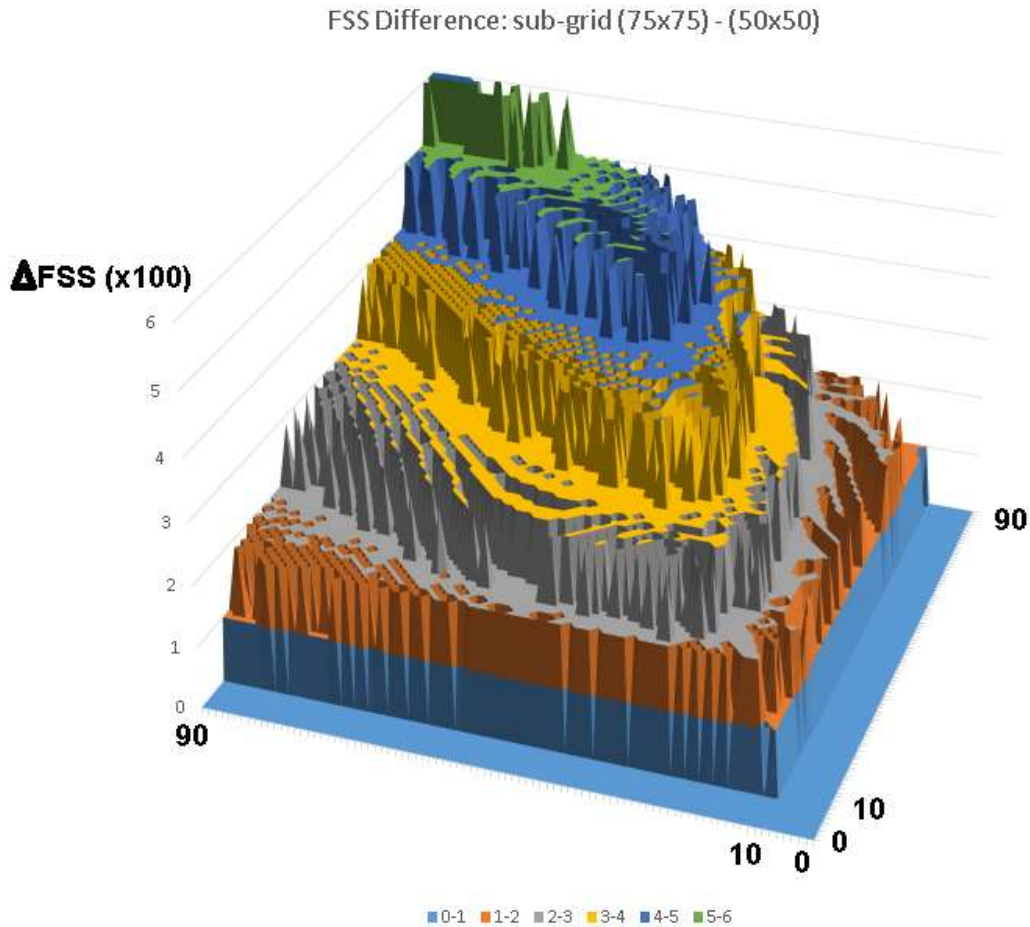


Figure 15: 3-D plot of the difference between the FSS for a 75x75-pixel and a 50x50-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradient respectively. The vertical axis displays the corresponding FSS difference.

2.4.4 The FSS for the 100x100-pixel sub-grid size

Further increasing the sub-grid window size by another 25 pixel in both directions (*Figure 16*) does not lead to a substantial improvement of the FSS ($\Delta FSS \leq 3$) (*Figure 18*). It also seems, taking into account the average time shift between ASII-TF data (12:00 UTC) and the related MetOp overpasses, that the horizontal shifts between both data sources are usually smaller than 100 pixel.

As for the previous sub-grid size, the FSS is well above 0.5 (= 50 in the 3-D plot) for similar hit-thresholds (red line in *Figure 17*).

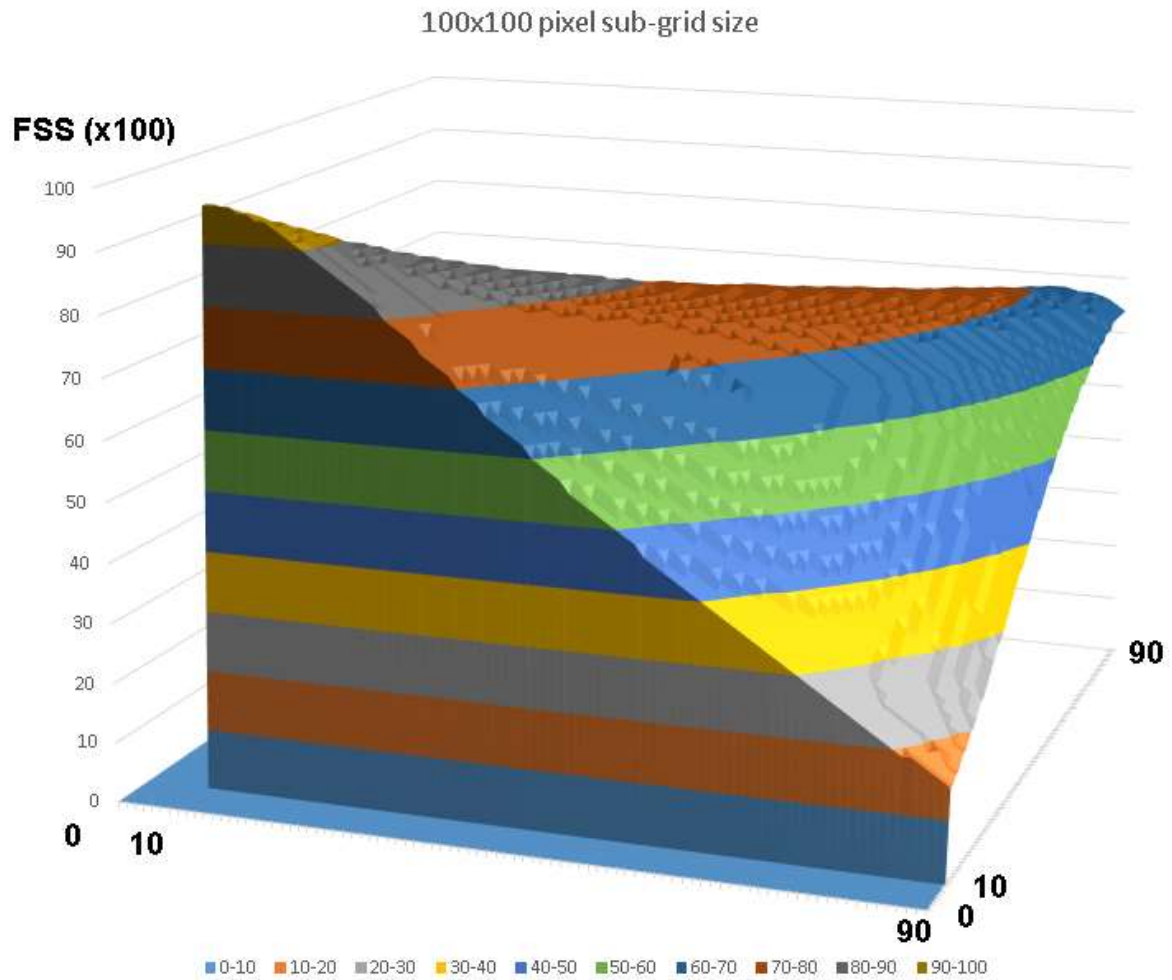


Figure 16: 3-D plot of the FSS for a 100x100-pixel sub-grid. The X-axis and the Y-axis display the hit-threshold applied to ASII-TF probabilities and to IASI L2 tropopause height gradients respectively. The vertical axis displays the corresponding FSS.

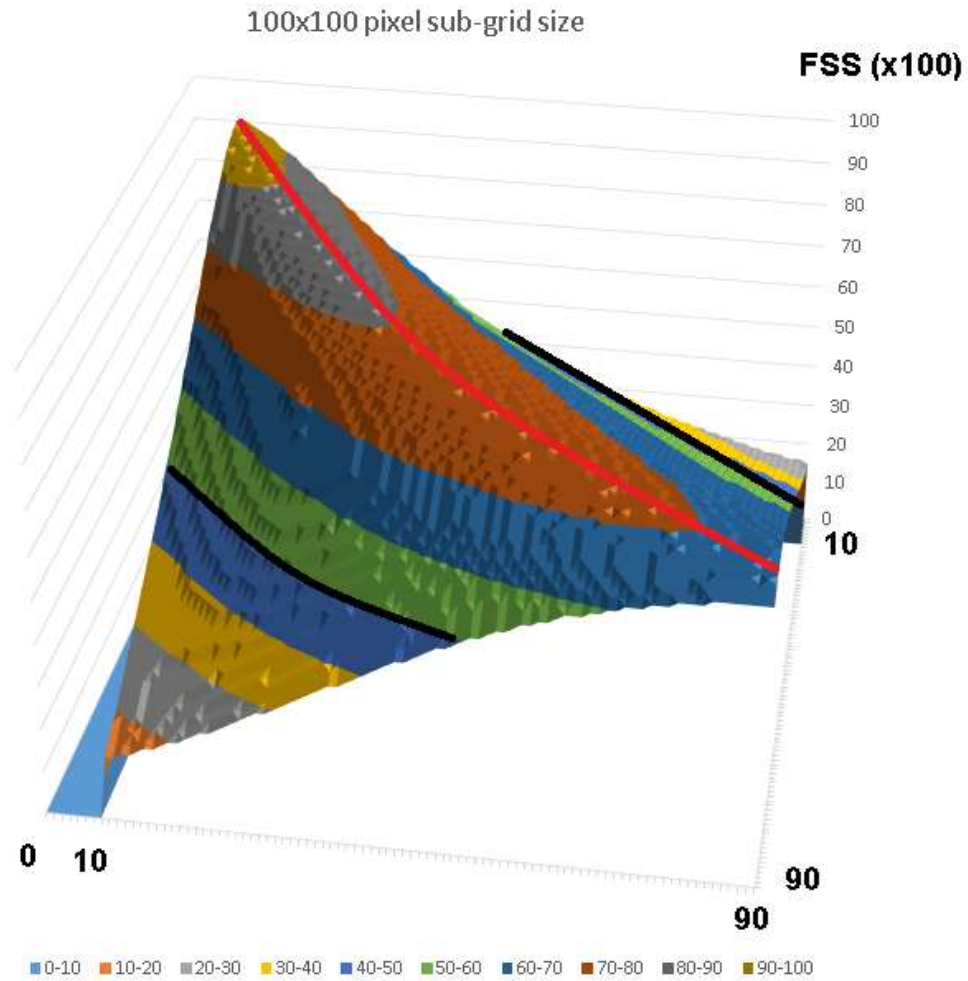


Figure 17: Same 3-D plot of the FSS for the 100x100-pixel sub-grid as in Figure 16, but seen from a different perspective. The red curve marks the ridge line of the 3-D plot and the black curve the FSS_{target} line.

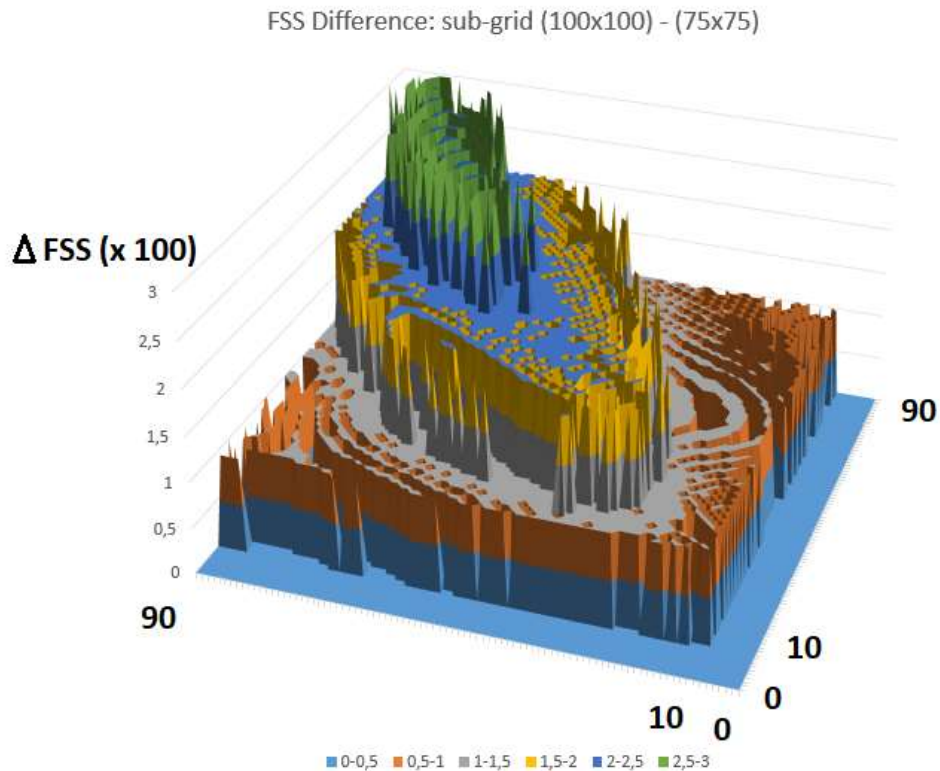


Figure 18: 3-D plot of the difference between the FSS with a 100x100-pixel and a 75x75-pixel sub-grid size. The X-axis and the Y-axis display the hit-thresholds applied to ASII-TF probabilities and to IASI L2 tropopause height gradient respectively. The colors display the corresponding FSS difference.

2.4.5 Some remarks on the behavior of the FSS with increasing hit-thresholds

The steady decrease of the FSS with increasing hit thresholds as shown in all the figures above is not a common feature when we look at individual cases. In approximately one third of the validation cases (in 36 out of 106 cases), we observe an upswing of the FSS with increasing hit-thresholds after an initial decrease. This behavior is exemplified in the validation case from 18 March 2021.

Contrary to the trend, **Figure 19** displays a validation dataset that shows higher FSS for stronger tropopause gradients (IASI L2) and higher probabilities (ASII-TF). From this, it follows that smaller areas (i.e., areas with high probability values in ASII-TF or with strong tropopause gradients in IASI L2) match better than the more extended areas with lower thresholds. Note that for very low thresholds, the FSS always show high values as the compared sub-grid areas are largely filled with hits. However, in that case the FSS_{target} threshold has to be raised well above 0.5 (i.e., 50 in the above plots), following Eq. 1.

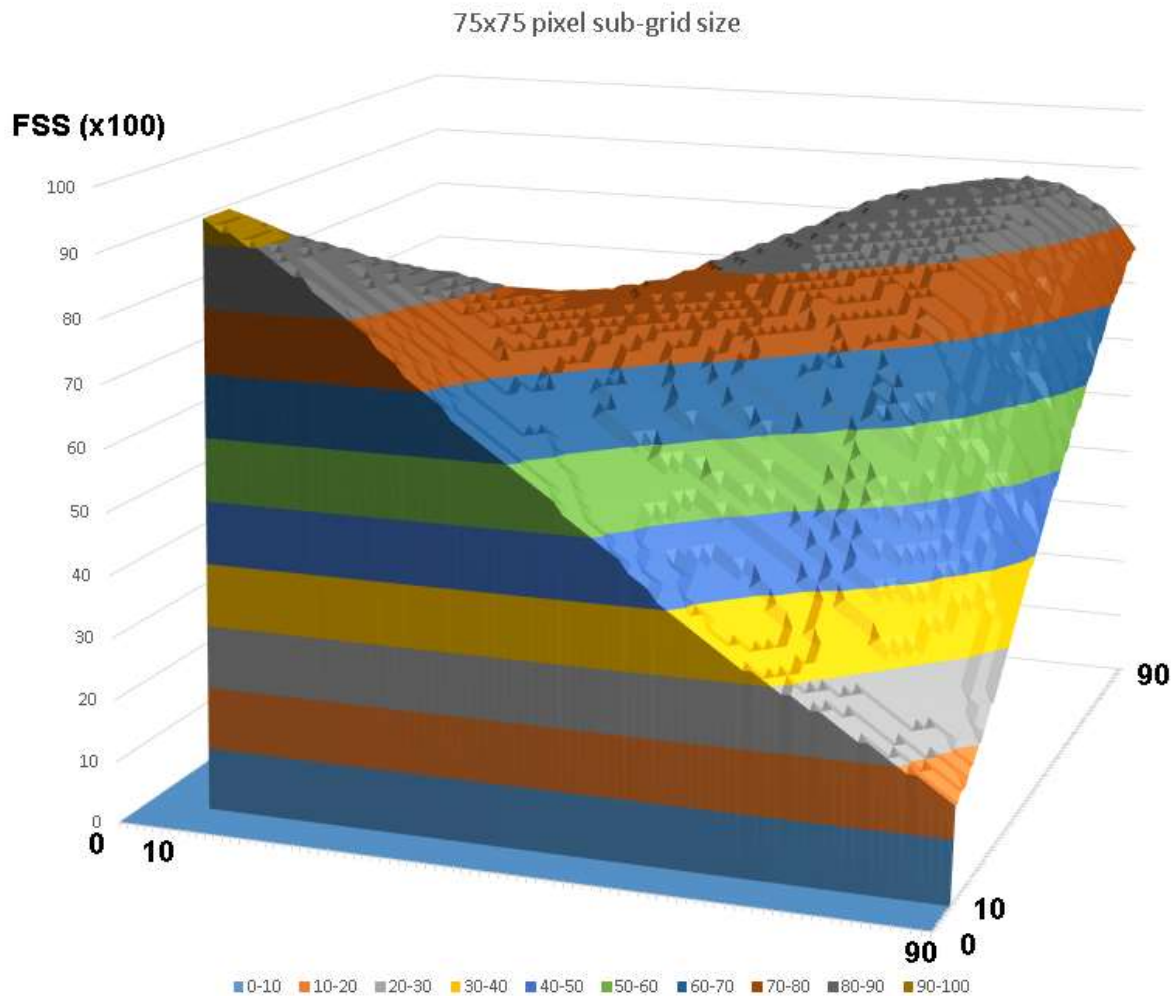


Figure 19: 3-D plot of the FSS for a 75x75-pixel sub-grid for 18 March 2021 at 12:00 UTC. The X-axis and the Y-axis display the ‘hit’ threshold applied to ASII-TF and IASI L2 tropopause height gradient respectively. The vertical axis displays the corresponding FSS.

2.5 SUMMARY

The ASII-TF product based on Himawari-8 data has been validated against tropopause gradients derived from IASI L2 sounding data. In a first step, a correlation analysis was performed, in a second step, the Fractions Skill Score was computed.

The correlation analysis shows positive values with nevertheless mediocre results (~0.44) that are mainly due to:

- Time shifts between the two data sources
- Different resolution of ASII-TF output and IASI L2 sounding data
- Different sensitivity regarding the derived parameter

To overcome the problems caused by the time shift and differing resolution of the used data sources, the FSS method was applied with varying sub-grid sizes and for different hit-thresholds.

The derived FSS were well above the FSS_{target} threshold as defined by Mittermaier and Roberts (2010) for all the validation cases and selected sub-grid areas when the corresponding hit-thresholds were chosen similar in magnitude for both data sources.

While in general the FSS drop with increasing hit-thresholds (i.e., with higher probability values in ASII-TF and stronger tropopause gradients in IASI L2), in about one third of the validation cases, this is not the case. Instead, high ASII-TF probability values showed better agreement than lower probability values there.

3. VALIDATION OF THE ASII-NG GRAVITY WAVE PRODUCT, V1.1

As described in the ATBD [RD.3], the ASII-NG product suite shall comprise modules relevant for aviation meteorology, through objective detection of patterns generally considered as indicative of (clear-air) turbulence. The idea behind the ASII-NG gravity wave product (ASII-GW) is to support the aviation forecaster in making his/her decision (from the meteorological material comprising much more than just the ASII-NG) by relieving him/her from a subjective search over the whole image by providing a reliable objective identification of the relevant pattern. In the case of ASII-GW, the relevant signal seen in Meteosat imagery are parallel stripes / ripples (see **Figure 20**), and – as concluded in the general discussion in the precursor validation report [RD.4] - the scope of ASII-GW validations is to give insight how well the outcome of the employed pattern recognition algorithm corresponds with a subjective diagnosis of these patterns.

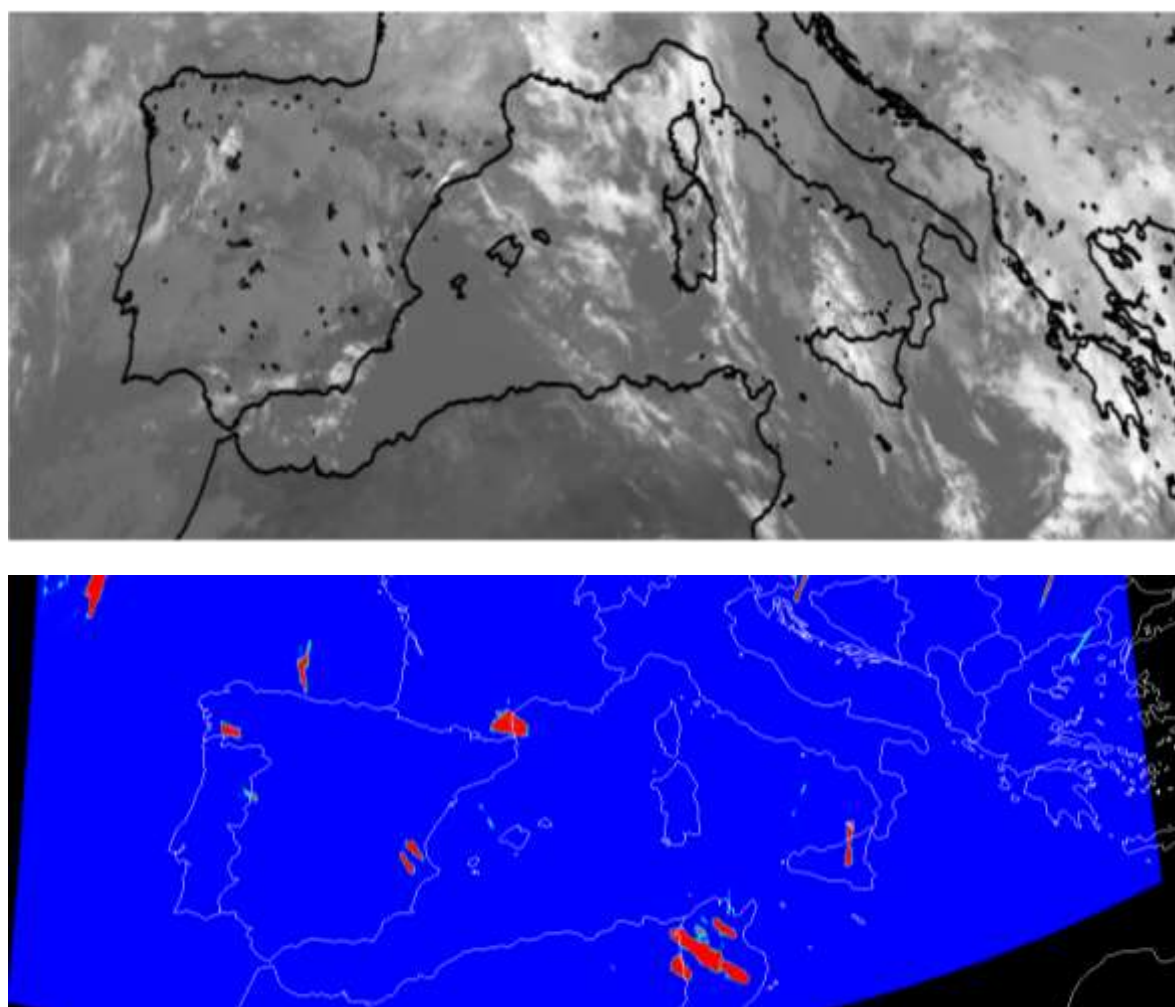






Figure 20: Case of 28 December 2021, 1200 UTC. Upper panel: MSG-4 IR10.8 image. Ripples characteristic for gravity waves and comparatively easily spotted in the image are present over Tunisia and the western part of the Gulf of Lion. Lower panel: The probability-of-occurrence of gravity waves, as analysed by the ASII-GW module, v2021, on the basis of the IR10.8 image (the colour table runs from dark blue = 0% to red=100%).

Since the last ASII-NG validation report [RD.4], the continuous interplay between development, tuning and validation has led to two peer-reviewed papers, constituting the documentation of the validation campaigns that led to v2021:

 NWC SAF	 ZAMG	Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG Issue: 2.0 Date: 10 January 2022 File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0 Page: 33/34
---	--	--	---

- The changes in the core algorithm between v2018 and v2021 are covered in Jann (2019). The ATBD [RD.3] also identifies the new algorithmic components as such. With the modified algorithm, it was possible to introduce an IR gravity wave analysis in v2021 (mentioned in [RD.4] as too contaminated by false alarms for the algorithm of v2018). **Figure 20** shows such an analysis.
- He et al. (2020) explored the application of the algorithm to the higher-resolution Himawari data and tuned algorithmic parameters for WV, IR and VIS bands (and moreover reconfirmed that the algorithmic change between v2018 and v2021 is a positive one also for this spatial resolution). The results for WV and IR went into v2021 for the analysis of Himawari and GOES-R imagery.

		Scientific and Validation Report for the “Automatic Satellite Image Interpretation – Next Generation” Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/ZAMG/SCI/VR/ASII-NG Issue: 2.0 Date: 10 January 2022 File: NWC-CDOP3-GEO-ZAMG-SCI-VR-ASII-NG_v2.0 Page: 34/34
---	---	--	---

4. REFERENCES

Faggian, N., B. Roux, P. Steinle and B. Ebert (2015): Fast calculation of the fractions skill score. *Mausam*, **66**, 457-466.

He, N., A. Jann and Y. Wang (2020): Objective detection of gravity waves in Himawari-8 imagery in support of aviation forecasting. *Meteorol. Z. (Contrib. Atm. Sci.)*, **29**, 323–332.

Jann, A. (2019): Objective detection of stripe patterns in satellite imagery caused by gravity waves: Lessons learnt from the southern hemisphere. *Trans. R. Soc. S. Afr.*, **74**, 163-172, <https://doi.org/10.1080/0035919X.2019.1596176>.

Mittermaier, M.P., and N. Roberts (2010): Inter-comparison of spatial forecast verification methods: identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354.

Mittermaier, M.P., N. Roberts and S.A. Thompson (2013): A long-term assessment of precipitation forecast skill using the Fractions Skill Score. *Meteorol. Appl.*, **20**, 176–186. DOI: 10.1002/met.296

Roberts, N.M., and H.W. Lean (2008): Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97.

Zhao, B., and B. Zhang (2018): Assessing hourly precipitation forecast skill with the Fractions Skill Score. *J. Meteor. Res.*, **32**, 135-145, <https://doi.org/10.1007/s13351-018-7058-1>.